BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

STIC-Biotech/ChemLib 5744

From: Nickol, Gary
Sent: Wednesday, January 02, 2002 2:03 PM
To: STIC-Biotech/ChemLib
Subject: 09/733956

1) Please search the amino acids of SEQ ID NO:2.

2) Please search DNA encoding SEQ ID NO:2.

Thanks,

CRF

Gary Nickol Ph.D.
AU:1642, Room 8D09, Mailbox 8E12
703-305-7143

	TYPE OF SEARCH:	VENDOR/COST(where applic.)
Searcher: _____	NA Sequences:_____	STN: _____
Phone: _____	AA Sequences:_____	DIALOG: _____
Location: _____	Structures:_____	Questel/Orbit: _____
Date Picked Up: _____	Bibliographic: _____	DRLink:_____
Date Completed: 1/9/02	Litigation: _____	Lexis/Nexis: _____
Searcher Prep/Review: _____	Full text:_____	Sequence Sys.: _____
Clerical: _____	Patent Family:_____	WWW/Internet: _____
Online time: _____	Other:_____	Other (specify): _____

GenCore version 4.5

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:16 ; Search time 12.66 Seconds
(without alignments)
719.892 Million cell updates/sec

Title: US-09-733-956-2
Perfect score: 2091
Sequence: 1 MNSTESNSSASVRPKVLSSI.....SQEASQTLLDSVYSHLPDLL 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2091	100.0	655	3	US-08-959-382-2	Sequence 2, Appli
2	1914	91.5	573	4	US-09-042-785A-2	Sequence 2, Appli
3	1527	73.0	605	4	US-09-042-785A-23	Sequence 23, Appl
4	431	20.6	84	4	US-09-042-785A-30	Sequence 30, Appl
5	431	20.6	84	4	US-09-042-785A-31	Sequence 31, Appl
6	115.5	5.5	2183	3	US-08-746-111-5	Sequence 5, Appli
7	99	4.7	2343	4	US-09-324-867-2	Sequence 2, Appli
8	95.5	4.6	287	4	US-08-985-950-6	Sequence 6, Appli
9	95.5	4.6	287	4	US-08-985-950-8	Sequence 8, Appli
10	93	4.4	465	3	US-08-788-231A-2	Sequence 2, Appli
11	91.5	4.4	1093	4	US-09-315-793-52	Sequence 52, Appl
12	90	4.3	1719	2	US-08-459-568-4	Sequence 4, Appli
13	90	4.3	1719	2	US-08-399-411-4	Sequence 4, Appli
14	90	4.3	1719	3	US-08-516-859A-4	Sequence 4, Appli
15	89	4.3	761	4	US-08-928-941D-1	Sequence 1, Appli
16	89	4.3	761	4	US-09-280-590A-1	Sequence 1, Appli
17	87.5	4.2	760	4	US-08-928-941D-29	Sequence 29, Appl
18	87.5	4.2	760	4	US-09-280-590A-29	Sequence 29, Appl
19	87	4.2	651	4	US-08-985-950-22	Sequence 22, Appl
20	86.5	4.1	895	4	US-08-827-962-19	Sequence 19, Appl
21	86.5	4.1	895	4	US-08-827-962-21	Sequence 21, Appl
22	86.5	4.1	1072	4	US-09-357-251-31	Sequence 31, Appl
23	86.5	4.1	1162	4	US-08-827-962-15	Sequence 15, Appl
24	86.5	4.1	1162	4	US-08-827-962-20	Sequence 20, Appl
25	86.5	4.1	1162	4	US-08-803-346-1	Sequence 1, Appli
26	86	4.1	1248	2	US-09-080-897-2	Sequence 2, Appli
27	86	4.1	1248	4	US-09-323-735-2	Sequence 2, Appli
28	86	4.1	1315	3	US-08-899-595-3	Sequence 3, Appli
29	85.5	4.1	339	1	US-08-431-080-18	Sequence 18, Appl
30	85.5	4.1	339	2	US-08-938-534-18	Sequence 18, Appl
31	85.5	4.1	1447	4	US-09-041-886-25	Sequence 25, Appl
32	85.5	4.1	1447	5	PCT-US94-05277-2	Sequence 2, Appli
33	85	4.1	405	4	US-09-413-574-2	Sequence 2, Appli
34	84.5	4.0	599	4	US-09-000-145-2	Sequence 2, Appli
35	84.5	4.0	646	4	US-09-625-188-10	Sequence 10, Appl
36	84	4.0	772	1	US-08-524-757-12	Sequence 12, Appl
37	84	4.0	2548	4	US-09-172-422-1	Sequence 1, Appli
38	83.5	4.0	1070	4	US-08-697-954-2	Sequence 2, Appli
39	83	4.0	1481	2	US-08-616-844-40	Sequence 40, Appl
40	83	4.0	1481	2	US-08-599-654-40	Sequence 40, Appl
41	83	4.0	1481	3	US-08-944-868A-40	Sequence 40, Appl
42	83	4.0	1481	3	US-08-944-423A-40	Sequence 40, Appl
43	83	4.0	1481	3	US-08-944-496-40	Sequence 40, Appl
44	82.5	3.9	771	3	US-09-121-964-9	Sequence 9, Appli
45	82.5	3.9	1230	2	US-08-968-542C-35	Sequence 35, Appl

ALIGNMENTS

```
RESULT   1
US-08-959-382-2
; Sequence 2, Application US/08959382
; Patent No. 6013476
;  GENERAL INFORMATION:
;    APPLICANT:  DEEN, KEITH
;    APPLICANT:  HURLE, MARK
;    APPLICANT:  YOUNG, PETER
;    APPLICANT:  TAN, K.B.
;    TITLE OF INVENTION:  Tumor Necrosis Related Receptor,
;    TITLE OF INVENTION:  TR7
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/959,382
;      FILING DATE:  28-OCT-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/041,796
;      FILING DATE:  02-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  655 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-959-382-2

  Query Match             100.0%;  Score 2091;  DB 3;  Length 655;
```

```
  Best Local Similarity   100.0%;  Pred. No. 4.6e-204;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 550

Qy     301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 610

Qy     361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     611 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 655


RESULT   2
US-09-042-785A-2
; Sequence 2, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  573 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
09-042-785A-2

Jan

  Query Match             91.5%;  Score 1914;  DB 4;  Length 573;
  Best Local Similarity   92.9%;  Pred. No. 3.8e-186;
  Matches  377;  Conservative    6;  Mismatches   21;  Indels    2;  Gaps    2;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||:|||:|||| || | |:||||||||||  |||  |:||||   | ||| |||||||
Db     169 MNSTDSNSTASVRTKVPSGIEEGTVPDNTSSTSGKEGTNRTLPNPPQVTHQQAPHHRHIL 228

Qy      61 KLLP-SMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCS 119
           |||| ||||| |||||| || ||||||||| |||||||||||||||||||:|||||
Db     229 KLLPSSMEAT-GEKSSTAIKAPKRGHPRQNAHKHFDINEHLPWMIVLFLLLVLVLIVVCS 287

Qy     120 IRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQ 179
           |||||||||||||||||||||||||:|||||||||| ||||||||||||||||||
Db     288 IRKSSRTLKKGPRQDPSAIVEKAGLKKSLTPTQNREKWIYYRNGHGIDILKLVAAQVGSQ 347

Qy     180 WKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND 239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     348 WKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND 407

Qy     240 VVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGF 299
           |||||||||||||||||||||||||||||||||||||| ||||| |||||||| |||| |
Db     408 VVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNVKLENSTLLTVEPSPLDKNKCF 467

Qy     300 FVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNP 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     468 FVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHILNP 527

Qy     360 EELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     528 EELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 573


RESULT   3
US-09-042-785A-23
; Sequence 23, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  605 amino acids
;      TYPE:  amino acid
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:16 ; Search time 12.66 Seconds
(without alignments)
719.892 Million cell updates/sec

Title: US-09-733-956-2
Perfect score: 2091
Sequence: 1 MNSTESNSSASVRPKVLSSI.....SQEASQTLLDSVYSHLPDLL 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 225032?2 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2091	100.0	655	3	US-08-959-382-2	Sequence 2, Appli
2	1914	91.5	573	4	US-09-042-785A-2	Sequence 2, Appli
3	1527	73.0	605	4	US-09-042-785A-23	Sequence 23, Appl
4	431	20.6	84	4	US-09-042-785A-30	Sequence 30, Appl
5	431	20.6	84	4	US-09-042-785A-31	Sequence 31, Appl
6	115.5	5.5	2183	3	US-08-746-111-5	Sequence 5, Appli
7	99	4.7	2343	4	US-09-324-867-2	Sequence 2, Appli
8	95.5	4.6	287	4	US-08-985-950-6	Sequence 6, Appli
9	95.5	4.6	287	4	US-08-985-950-8	Sequence 8, Appli
10	93	4.4	465	3	US-08-788-231A-2	Sequence 2, Appli
11	91.5	4.4	1093	4	US-09-315-793-52	Sequence 52, Appl
12	90	4.3	1719	2	US-08-459-568-4	Sequence 4, Appli
13	90	4.3	1719	2	US-08-399-411-4	Sequence 4, Appli
14	90	4.3	1719	3	US-08-516-859A-4	Sequence 4, Appli
15	89	4.3	761	4	US-08-928-941D-1	Sequence 1, Appli
16	89	4.3	761	4	US-09-280-590A-1	Sequence 1, Appli
17	87.5	4.2	760	4	US-08-928-941D-29	Sequence 29, Appl
18	87.5	4.2	760	4	US-09-280-590A-29	Sequence 29, Appl
19	87	4.2	651	4	US-08-985-950-22	Sequence 22, Appl
20	86.5	4.1	895	4	US-08-827-962-19	Sequence 19, Appl
21	86.5	4.1	895	4	US-08-827-962-21	Sequence 21, Appl
22	86.5	4.1	1072	4	US-09-357-251-31	Sequence 31, Appl
23	86.5	4.1	1162	4	US-08-827-962-15	Sequence 15, Appl
24	86.5	4.1	1162	4	US-08-827-962-20	Sequence 20, Appl
25	86.5	4.1	1162	4	US-08-803-346-1	Sequence 1, Appli
26	86	4.1	1248	2	US-09-080-897-2	Sequence 2, Appli
27	86	4.1	1248	4	US-09-323-735-2	Sequence 2, Appli
28	86	4.1	1315	3	US-08-899-595-3	Sequence 3, Appli
29	85.5	4.1	339	1	US-08-431-080-18	Sequence 18, Appl
30	85.5	4.1	339	2	US-08-938-534-18	Sequence 18, Appl
31	85.5	4.1	1447	4	US-09-041-886-25	Sequence 25, Appl
32	85.5	4.1	1447	5	PCT-US94-05277-2	Sequence 2, Appli
33	85	4.1	405	4	US-09-413-574-2	Sequence 2, Appli
34	84.5	4.0	599	4	US-09-000-145-2	Sequence 2, Appli
35	84.5	4.0	646	4	US-09-625-188-10	Sequence 10, Appl
36	84	4.0	772	1	US-08-524-757-12	Sequence 12, Appl
37	84	4.0	2548	4	US-09-172-422-1	Sequence 1, Appli
38	83.5	4.0	1070	4	US-08-697-954-2	Sequence 2, Appli
39	83	4.0	1481	2	US-08-616-844-40	Sequence 40, Appl
40	83	4.0	1481	2	US-08-599-654-40	Sequence 40, Appl
41	83	4.0	1481	3	US-08-944-868A-40	Sequence 40, Appl
42	83	4.0	1481	3	US-08-944-423A-40	Sequence 40, Appl
43	83	4.0	1481	3	US-08-944-496-40	Sequence 40, Appl
44	82.5	3.9	771	3	US-09-121-964-9	Sequence 9, Appli
45	82.5	3.9	1230	2	US-08-968-542C-35	Sequence 35, Appl

ALIGNMENTS

```
RESULT   1
US-08-959-382-2
; Sequence 2, Application US/08959382
; Patent No. 6013476
;  GENERAL INFORMATION:
;    APPLICANT:  DEEN, KEITH
;    APPLICANT:  HURLE, MARK
;    APPLICANT:  YOUNG, PETER
;    APPLICANT:  TAN, K.B.
;    TITLE OF INVENTION:  Tumor Necrosis Related Receptor,
;    TITLE OF INVENTION:  TR7
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/959,382
;      FILING DATE:  28-OCT-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/041,796
;      FILING DATE:  02-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  655 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-959-382-2


  Query Match             100.0%;  Score 2091;  DB 3;  Length 655;
```

```
  Best Local Similarity   100.0%;  Pred. No. 4.6e-204;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 550

Qy     301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 610

Qy     361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     611 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 655


RESULT   2
US-09-042-785A-2
; Sequence 2, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  573 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-042-785A-2


  Query Match           91.5%;  Score 1914;  DB 4;  Length 573;
  Best Local Similarity   92.9%;  Pred. No. 3.8e-186;
  Matches  377;  Conservative    6;  Mismatches   21;  Indels    2;  Gaps    2;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||:|||:|||| || | |:||||||||||  |||  |:||||   | ||| |||||||
Db     169 MNSTDSNSTASVRTKVPSGIEEGTVPDNTSSTSGKEGTNRTLPNPPQVTHQQAPHHRHIL 228

Qy      61 KLLP-SMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCS 119
           |||| ||||| |||||| || |||||||| ||||||||||||||||||||:|||||
Db     229 KLLPSSMEAT-GEKSSTAIKAPKRGHPRQNAHKHFDINEHLPWMIVLFLLLVLVLIVVCS 287

Qy     120 IRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQ 179
           |||||||||||||||||||||||||||:||||||||||| ||||||||||||||||||
Db     288 IRKSSRTLKKGPRQDPSAIVEKAGLKKSLTPTQNREKWIYYRNGHGIDILKLVAAQVGSQ 347

Qy     180 WKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND 239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     348 WKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND 407

Qy     240 VVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGF 299
           ||||||||||||||||||||||||||||||||||||| ||||| |||||||| |||| |
Db     408 VVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNVKLENSTLLTVEPSPLDKNKCF 467

Qy     300 FVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNP 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     468 FVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHILNP 527

Qy     360 EELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     528 EELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 573


RESULT   3
US-09-042-785A-23
; Sequence 23, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  605 amino acids
;      TYPE:  amino acid
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-042-785A-23


  Query Match             73.0%;  Score 1527;  DB 4;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 1e-146;
  Matches  294;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQD 294
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQD 544


RESULT   4
US-09-042-785A-30
; Sequence 30, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-09-042-785A-30


  Query Match             20.6%;  Score 431;  DB 4;  Length 84;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-36;
  Matches   84;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     165 GIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEAS 224
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEAS 60

Qy     225 LAQLISALRQHRRNDVVEKIRGLM 248
           |||||||||||||||||||||||||
Db      61 LAQLISALRQHRRNDVVEKIRGLM 84


RESULT   5
US-09-042-785A-31
; Sequence 31, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-09-042-785A-31


  Query Match             20.6%;  Score 431;  DB 4;  Length 84;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-36;
  Matches   84;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     165 GIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEAS 224
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEAS 60

Qy     225 LAQLISALRQHRRNDVVEKIRGLM 248
           |||||||||||||||||||||||||
Db      61 LAQLISALRQHRRNDVVEKIRGLM 84


RESULT   6
US-08-746-111-5
```

```
; Sequence 5, Application US/08746111
; Patent No. 6066778
;  GENERAL INFORMATION:
;    APPLICANT:  Ginsburg, David
;    APPLICANT:  Cui, Jisong
;    TITLE OF INVENTION:  Compositions And Methods For Screening
;    TITLE OF INVENTION:  Compounds For Anticoagulant Activity
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/746,111
;      FILING DATE:  06-NOV-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  UM-02536
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2183 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-746-111-5


  Query Match             5.5%;  Score 115.5;  DB 3;  Length 2183;
  Best Local Similarity   19.9%;  Pred. No. 0.042;
  Matches   97;  Conservative   49;  Mismatches  197;  Indels  145;  Gaps   20;

Qy       5 ESNSSASVRPKVLSSIQEGTVPD--NTSSARGKEDVNKTLPNL----------QVVNHQQ 52
           :|||      ::|| |     : |    |     |        | ||              |:
Db     773 DSNSS-----RILSKIINNNLKDFQRTLPGSGATVAGTLLRNLIGLDENFVLNSSTEHRS 827

Qy      53 GPHHRH----------ILKLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHF-------D 95
             :| :          :: |||     | :     |  | | |    |:
Db     828 SSYHENDMENPQSNITMVYLLPLGPKGSGNREQDKPKTIKTGRPHMMKHRFSWMKAPAGK 887

Qy      96 INEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNRE 155
             |                  |  |:| :  |     || ::    ||: :|
Db     888 TGRH------------------SNPKNSYSGMKSEEDIPSELIP---LKQKITSKFLNR 925

Qy     156 KW----------IYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTAD 205
           :|          |   || ' |: ||                |: : :      | :
Db     926 RWRVASEKGSYEIIAANGEDTDVDKLT---------------NSPQNQNITVPRGESTS 969

Qy     206 HERAYAALQHWTIRGPE-----ASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDKLA 260
           |          | | |      : :   :||   |   :| :  :    : :  |||
Db     970 HTN--------TTRKPSDLPTFSGVGHKSPHVRQEEENSGFQKRQLFIRTRKKKKNKKLA 1021

Qy     261 L--PMSPSPLSP----SPIPSPNAKLENSALL-------TVEPSPQDKNKGFFVDESEP- 306
           |  |:||    |     :  | |: :| | :||        : |      :      | | |
Db    1022 LHSPLSPRGFDPLRGHNHSPFPDRRLLNHSLLLHKSNETALSPDLNQTSPSMSTDRSLPD 1081

Qy     307 --LLRCDSTSSGSSALSRNGSFITKEKKDTVLRQ------------VRLDPCDLQPIFD- 351
               : |   ||:| |  . :|  |   |              | | :|    |
Db    1082 YNQYSKNDTEQMSSSLDLYQSVPAEEHSPTFPAQDPDQTHSTTDPSYRSSPPELSQGLDY 1141

Qy     352 DMLHFLNPEELRVIEEIPQAEDKLD-------------RLFEIIGVKSQEASQTLLDSV 397
```

```
           |: |   |::: :    |  |                 || |    | |  ||:   :
Db    1142 DLSHDFYPDDIGLTSFFPDQSQKSSFSSDDDQAIPSSDLSLFTI----SPELDQTI---I 1194

Qy     398 YSHLPDLL 405
           |  |  ||
Db    1195 YPDLDQLL 1202


RESULT   7
US-09-324-867-2
; Sequence 2, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:
;  APPLICANT: Lillicrap, David
;  APPLICANT: Cameron, Cherie
;  APPLICANT: No. 6251632ley, Colleen
;  APPLICANT: Horrocks, L. Suzanne Hoyle
;  APPLICANT: Hough, Christine
;  TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
;  FILE REFERENCE: 1669.0010002/JAG/BJD
;  CURRENT APPLICATION NUMBER: US/09/324,867A
;  CURRENT FILING DATE: 1999-06-03
;  EARLIER APPLICATION NUMBER: 09/035,141
;  EARLIER FILING DATE: 1998-03-059
;  EARLIER APPLICATION NUMBER: 60/039,953
;  EARLIER FILING DATE: 1997-03-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2343
;   TYPE: PRT
;   ORGANISM: Canis familiaris
US-09-324-867-2


  Query Match             4.7%;  Score 99;  DB 4;  Length 2343;
  Best Local Similarity   19.4%;  Pred. No. 2.2;
  Matches   89;  Conservative   62;  Mismatches  153;  Indels  154;  Gaps   19;

Qy      10 ASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQV-VNHQQGPHHRHILK------- 61
           ||:||::  |      |:                | ||: :|   | :    ||
Db     846 ASLRPELRHSEDREFTPE--------------PELQLRLNENLGTNTTVELKKLDLKIS 890

Qy      62 -----LLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIV---------- 105
                |: |      : ::   |    | |  ::|     | ||  ::
Db     891 SSSDSLMTSPTIPSDKLAAATEKTGSLGPPNMSVH----FNSHLGTIVFGNNSSHLIQSG 946

Qy     106 ------------LFLLLVLVVIVVCSIR------KSSRTLKKGPRQDPSAIVEK------ 141
                       |   |: |   |:|      :|:|  |:   : |:::::
Db     947 VPLELSEEDNDSKLLEAPLMNIQESSLRENVLSMESNRLFKEERIRGPASLIKDNALFKV 1006

Qy     142 --AGLKKSMTP---TQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVA 196
             : :| :  |   | ||:          : |  |:     | |:||     |   :||
Db    1007 NISSVKTNRAPVNLTTNRKT--------RVAIPTLLIENSTSVWQDI-MLERNTEFKEVT 1057

Qy     197 AFSNGYT-ADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDTTQLE 255
           :  :  | |        | | :    : :|:: :    | :             ||   |
Db    1058 SLIHNETFMDRNTTALGLNHVS---NKTTLSKNVEMAHQKK------------EDPVPLR 1102

Qy     256 TDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLLRCDSTSS 315
           :  |  |  |  |  | :      :||      |||:                    ||
Db    1103 AENPDLSSSKIPFLPDWIKTHG---KNSLSSEQRPSPK----------------QLTSL 1142

Qy     316 GSSALSRNGSFITKEK---------KDTVLRQVRLDPCDLQPIF--DDMLHFLN------ 358
           ||    :: :|:::||          |||          :|| ||  :  : | |
Db    1143 GSEKSVKDQNFLSEEKVVVGEDEFTKDT----------ELQEIFPNNKSIFFANLANVQE 1192

Qy     359 ----------PEELRVIEEIPQAEDKLDRLFEIIGVKS 386
                    |||:   |:: |    | :   :|| |:
Db    1193 NDTYNQEKKSPEEIERKEKLTQENVALPQAHTMIGTKN 1230


RESULT   8
```

```
US-08-985-950-6
; Sequence 6, Application US/08985950
; Patent No. 6140076
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  287 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-985-950-6


  Query Match             4.6%;  Score 95.5;  DB 4;  Length 287;
  Best Local Similarity   21.6%;  Pred. No. 0.16;
  Matches   50;  Conservative   32;  Mismatches   75;  Indels   75;  Gaps   12;

Qy      27 DNTSSARGKEDVNKTLPNL--------QVVNHQQGPH------------HRHILKLLPSM 66
           :: |:    |||::  |:          |     ||:               |:||
Db      63 ESRSTYNDTEDVSQASPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSEQSDYLELLVK- 121

Qy      67 EATGGEKSSTPIKGPKRG---HPRQNLHKHFDINEHLP--------------WMIVLFLL 109
           | :||  |     |   |    |  | |     ||| |                 : |:||
Db     122 ETSGGPDSPDTEPGSSAGPTQRPSDNSH-----NEHAPASQGLKAEHLYILIGVSVVFLF 176

Qy     110 LVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDIL 169
           :|:::: |  |::   :|:|| :            :   | |  :          :|:|
Db     177 CLLLLVLFCLHRQNQ--IKQGPPRSKD---------EEQKPQQRPDL--------AVDVL 217

Qy     170 KLVAAQV---GSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTI 218
           :  | :    |  ||         | : :| : |  :  |  || | || :
Db     218 ERTADKATVNGLPEKD--------RETDTSALAAG--SSQEVTYAQLDHWAL 259


RESULT   9
US-08-985-950-8
; Sequence 8, Application US/08985950
; Patent No. 6140076
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  287 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-985-950-8


  Query Match             4.6%;  Score 95.5;  DB 4;  Length 287;
  Best Local Similarity   21.6%;  Pred. No. 0.16;
  Matches   50;  Conservative   32;  Mismatches   75;  Indels   75;  Gaps   12;

Qy      27 DNTSSARGKEDVNKTLPNL--------QVVNHQQGPH------------HRHILKLLPSM 66
           :: |:    |||::  |:          |     ||:               |:||
Db      63 ESRSTYNDTEDVSQASPSESEARFRIDSVSEGNAGPYRCIYYKPPKWSEQSDYLELLVK- 121

Qy      67 EATGGEKSSTPIKGPKRG---HPRQNLHKHFDINEHLP--------------WMIVLFLL 109
           | :||  |     |   |    |  | |     ||| |                 : |:||
Db     122 ETSGGPDSPDTEPGSSAGPTQRPSDNSH-----NEHAPASQGLKAEHLYILIGVSVVFLF 176

Qy     110 LVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDIL 169
           :|:::: |  |::   :|:|| :            :   | |  :          :|:|
Db     177 CLLLLVLFCLHRQNQ--IKQGPPRSKD---------EEQKPQQRPDL--------AVDVL 217

Qy     170 KLVAAQV---GSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTI 218
           :  | :    |  ||         | : :| : |  :  |  || | || :
Db     218 ERTADKATVNGLPEKD--------RETDTSALAAG--SSQEVTYAQLDHWAL 259


RESULT  10
US-08-788-231A-2
; Sequence 2, Application US/08788231A
; Patent No. 6019974
;  GENERAL INFORMATION:
;    APPLICANT:  L'Hernault, Steven W.
;    TITLE OF INVENTION:  SPE-4 RELATED PEPTIDES, ANTIBODIES AND
```

```
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  US
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/788,231A
;      FILING DATE:  24-JAN-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/010,672
;      FILING DATE:  26-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  60-95
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 499-8080
;      TELEFAX:  (303) 499-8089
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  465 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
US-08-788-231A-2


  Query Match             4.4%;  Score 93;  DB 3;  Length 465;
  Best Local Similarity   23.7%;  Pred. No. 0.65;
  Matches   47;  Conservative   33;  Mismatches   62;  Indels   56;  Gaps   10;

Qy    169 LKLVAAQVGSQWKDIYQFLC-NASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQ 227
          || |  :     | :   :  :|:|: : | ||             |  |  | |::: :
Db    211 LKKVQEKASDYSKCVLNLIMFSANEKRLTAGSN-----------QEETNEGEESTIRR 257

Qy    228 LIS------ALRQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLE 281
          :        |: : ::  :|||                  : ::|  :|:     |
Db    258 TVKQTIEYYTKREAQDDEFYQKIR----------------QRRAAINPDSVPT-----E 295

Qy    282 NSALLTVEPSP---QDKN--KGFFVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVL 336
          :| |:  ||||   ::||  :    |||:       |||||| || : |  |     |
Db    296 HSPLVEAEPSPIELKEKNSTEELSDDESD----TSETSSGSSNLSSSDSSTTVSTSDI-- 349

Qy    337 RQVRLDPCDLQPIFDDML 354
            : || |  :||::
Db    350 --STAEECD-QKEWDDLV 364


RESULT  11
US-09-315-793-52
; Sequence 52, Application US/09315793
; Patent No. 6221597
; GENERAL INFORMATION:
;  APPLICANT: Roberts, Christopher J.
;  TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
;  TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION
;  TITLE OF INVENTION: DRUGS
;  FILE REFERENCE: 9301-048
;  CURRENT APPLICATION NUMBER: US/09/315,793
;  CURRENT FILING DATE: 1999-05-21
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 1093
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-315-793-52


  Query Match             4.4%;  Score 91.5;  DB 4;  Length 1093;
  Best Local Similarity   22.4%;  Pred. No. 3.7;
  Matches   73;  Conservative   51;  Mismatches  119;  Indels   83;  Gaps   17;

Qy    142 AGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGS----------QWKDIYQ--FLCN 189
          | :|| :||  |  | ::    ||  :: :  :  ||          :  : ||   : |
Db    595 AQIKKLITPRPN-GKILFKRIIHGNRLVDIKQSAYGSKQVFPTDVSIKQTNFYQGSIMSN 653

Qy    190 AS----EREVAAFSNGY-----TAD----------HERAYAALQH--------------- 215
              | |:    | |     | |          || :  | ::
Db    654 EQKIRIENEIINLKNEYNDRKSTLDALSNQKSGYRHELSELASKNDDINREAHQLNEIRK 713

Qy    216 -WTIR-GPEASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPI 273
           :|:|     :| : :  |::  | || :||: : :   ||   :  | :|    |   :
Db    714 KYTMRKSTIETLREKLDQLKREARKDVSQKIKDIDDQIQQLLLKQRHL-LSKMASSMKSL 772

Qy    274 PSPNAKLENSALLTVEPSPQDKNK----GFFVDESEPLLRCDSTSSGSSALSRNGSFITK 329
           :   :| :: :|  |    | :     ||| :   |          |       |: |
Db    773 KNCQKELISTQILQFEAQNMDVSMNDVIGFFNEREADL---------KSQYEDKKKFV-K 822

Qy    330 EKKDTV-----LRQVRLDPCDLQPIFD-DMLHFLNPEELRVIEE----IPQAEDKLDRL- 378
          | :||      :|::|         :| |    ||    :  ||    :   :| ||:|
Db    823 EMRDTPEFQSWMREIR--------SYDQDTKEKLNKVAEKYEEEGNFNLSFVQDVLDKLE 874

Qy    379 FEIIGVKSQEASQTLLDSVYSHLPDL 404
            ||  |   |:: |:|| | : | :|
Db    875 SEIAMVNHDESAVTILDQVTAELREL 900


RESULT  12
US-08-459-568-4
; Sequence 4, Application US/08459568
; Patent No. 5811304
;  GENERAL INFORMATION:
;    APPLICANT:  Huang, Shi
;    TITLE OF INVENTION:  Retinoblastoma Protein - Interacting
;    TITLE OF INVENTION:  Zinc Finger Proteins
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,568
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,411
;      FILING DATE:  06-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LJ 1264
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  4:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-459-568-4


  Query Match             4.3%;  Score 90;  DB 2;  Length 1719;
  Best Local Similarity   25.4%;  Pred. No. 11;
  Matches   35;  Conservative   15;  Mismatches   50;  Indels   38;  Gaps    5;

Qy     261 LPMSPSPLSPSPIPSP--NAKLENSALL---TVEPSPQDKNKGFFVDESEPLLRCDS--- 312
           ||   |  || | |||  ||  ::   :   || |||        :   |||:   |
Db     996 LPAPSSSASPHPCPSPLSNATAQSPLPILSPTVSPSPSP------IPPVEPLMSAASPGP 1049

Qy     313 --TSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ 370
             || ||: | : || :     :                        :|  |  |  :
Db    1050 PTLSSSSSSSSSSSFSSSSSSSP--------------------SPPPLSAISSVVS 1087

Qy     371 AEDKLDRLFEIIGVKSQE 388
           : | |:    :|  | :|
Db    1088 SGDNLEASLPMISFKQEE 1105


RESULT  13
US-08-399-411-4
; Sequence 4, Application US/08399411
; Patent No. 5831008
;  GENERAL INFORMATION:
;    APPLICANT:  Huang, Shi
;    TITLE OF INVENTION:  Retinoblastoma Protein - Interacting
;    TITLE OF INVENTION:  Zinc Finger Proteins
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,411
;      FILING DATE:  06-MAR-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LJ 1264
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-399-411-4


  Query Match             4.3%;  Score 90;  DB 2;  Length 1719;
  Best Local Similarity   25.4%;  Pred. No. 11;
  Matches   35;  Conservative   15;  Mismatches   50;  Indels   38;  Gaps    5;

Qy     261 LPMSPSPLSPSPIPSP--NAKLENSALL---TVEPSPQDKNKGFFVDESEPLLRCDS--- 312
           ||   |  || | |||  ||  ::   :   || |||        :   |||:   |
Db     996 LPAPSSSASPHPCPSPLSNATAQSPLPILSPTVSPSPSP------IPPVEPLMSAASPGP 1049

Qy     313 --TSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ 370
             || ||: | : || :     :                        :|  |  |  :
Db    1050 PTLSSSSSSSSSSSFSSSSSSSP--------------------SPPPLSAISSVVS 1087

Qy     371 AEDKLDRLFEIIGVKSQE 388
           : | |:    :|  | :|
Db    1088 SGDNLEASLPMISFKQEE 1105


RESULT  14
US-08-516-859A-4
; Sequence 4, Application US/08516859A
; Patent No. 6069231
;  GENERAL INFORMATION:
;    APPLICANT:  Huang, Shi
;    TITLE OF INVENTION:  Retinoblastoma Protein - Interacting
;    TITLE OF INVENTION:  Zinc Finger Proteins
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/516,859A
;      FILING DATE:  18-AUG-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/399,411
;      FILING DATE:  06-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/292,683
;      FILING DATE:  18-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LJ 1776
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-516-859A-4


  Query Match             4.3%;  Score 90;  DB 3;  Length 1719;
  Best Local Similarity   25.4%;  Pred. No. 11;
  Matches   35;  Conservative   15;  Mismatches   50;  Indels   38;  Gaps    5;

Qy     261 LPMSPSPLSPSPIPSP--NAKLENSALL---TVEPSPQDKNKGFFVDESEPLLRCDS--- 312
           ||   |  || | |||  ||  ::   :   || |||        :   |||:   |
Db     996 LPAPSSSASPHPCPSPLSNATAQSPLPILSPTVSPSPSP------IPPVEPLMSAASPGP 1049

Qy     313 --TSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ 370
             || ||: | : || :     :                        :|  |  |  :
Db    1050 PTLSSSSSSSSSSSFSSSSSSSP--------------------SPPPLSAISSVVS 1087

Qy     371 AEDKLDRLFEIIGVKSQE 388
           : | |:    :|  | :|
Db    1088 SGDNLEASLPMISFKQEE 1105
```

```
RESULT 15
US-08-928-941D-1
; Sequence 1, Application US/08928941D
; Patent No. 6180763
;  GENERAL INFORMATION:
;    APPLICANT:  Hirai, Hiroshi
;    APPLICANT:  Sherr, Charles
;    TITLE OF INVENTION:  CYCLIN-D BINDING FACTOR, AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David A. Jackson, Esq.
;      STREET:  411 Hackensack Ave, Continental Plaza, 4th
;      STREET:  Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,941D
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  1340-1-002 N CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  761 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Mus musculus
US-08-928-941D-1


  Query Match             4.3%;  Score 89;  DB 4;  Length 761;
  Best Local Similarity   20.6%;  Pred. No. 3.7;
  Matches   72;  Conservative   38;  Mismatches  113;  Indels  126;  Gaps   17;

Qy    136 SAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW--KDIYQFLCNASER 193
          :|: |: | :   :  | | ||: | |             |::|  :|   :   :|
Db    307 AAVAERVGTR---SEKQCRSKWLNYLNW---------KQSGGTEWTKEDEINLILRIAEL 354

Qy    194 EVA--------AFSNGYTADHERAYAALQHWTIR-----GPEASLAQLISALRQHRRND- 239
          :||         : |:::     :   : |||:       : |   ||  |:|   |
Db    355 DVADENDINWDLLAEGWSSVRSPQWLRSKWWTIKRQIANHKDVSFPVLIKGLKQLHENQK 414

Qy    240 ----VVEKIRG-------------------LMEDTTQLETDKLALPMSPSPLSPSPIP- 274
              ::|   |                  :|| | :         ||||::   ||
Db    415 NNPVLLENKSGSGVPNSNCNSSVQHVQIRVARLEDNTAI---------SPSPMAALQIPV 465

Qy    275 ---------SPNAK-------LENSALLTVE-----PSPQDKNKGFFVDESEPLLRCDST 313
                  || |         | :  | | |      |        | :       ||:  |:
Db    466 QITHVSSTDSPAASADSETITLNSGTLQTFEILPSFPLQPTGTPGTY------LLQTSSS 519

Qy    314 SSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEEL---------- 362
              |: |         |      |   : |   ::| |:|| | .
Db    520 QGLPLTLTTN----------PTLTLAAAAPASPEQI---IVHALSPEHLLNTSDNVTVQC 566

Qy    363 ---RVIEEIPQAEDKLDRLFEIIGVKSQE---ASQTLLDSVYSHLPDLL 405
              ||| :    ||       |   |||       |  | :   || |
Db    567 HTPRVIIQTVATED-------ITSSLSQEELTVDSDLHSSDFPEPPDAL 608


Search completed: January  4, 2002, 08:38:13
Job time: 117 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:11 ; Search time 24.92 Seconds
(without alignments)
1203.840 Million cell updates/sec

Title: US-09-733-956-2
Perfect score: 2091
Sequence: 1 MNSTESNSSASVRPKVLSSI.....SQEASQTLLDSVYSHLPDLL 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2091	100.0	405	21	AAY67947	Human DETH protein
2	2091	100.0	631	20	AAY05678	Human full-length
3	2091	100.0	655	19	AAW75792	Human tumour necro
4	2091	100.0	655	20	AAY41693	Human PRO868 prote
5	2091	100.0	655	20	AAW81059	Novel human tumor
6	2091	100.0	655	21	AAB26981	Human tumour necro
7	2091	100.0	655	21	AAB01349	Tumour necrosis fa
8	2091	100.0	655	21	AAB01338	DR-6 death recepto
9	2091	100.0	655	21	AAY77460	Human TNF receptor
10	2091	100.0	691	21	AAY92846	Human osteoprotegr
11	2089	99.9	655	21	AAB44249	Human PRO868 (UNQ4
12	2087	99.8	655	22	AAB93023	Human protein sequ
13	1914	91.5	573	20	AAY05695	Mouse TNF receptor
14	1527	73.0	605	20	AAY05697	Human TNF receptor
15	1396	66.8	600	21	AAY92845	Murine osteoproteg
16	1345	64.3	393	20	AAY28450	A human tumour nec
17	1322	63.2	508	20	AAY05680	Tumour necrosis fa
18	868	41.5	444	21	AAY77461	Human TNF receptor
19	439	21.0	85	22	AAM15484	Peptide #1918 enco
20	439	21.0	85	22	AAM16947	Peptide #3381 enco
21	439	21.0	85	22	AAM27972	Peptide #2009 enco
22	439	21.0	85	22	AAM29434	Peptide #3471 enco
23	439	21.0	85	22	AAM03241	Peptide #1923 enco
24	439	21.0	85	22	AAM04642	Peptide #3324 enco
25	431	20.6	84	20	AAY05698	Human TNF receptor
26	431	20.6	84	20	AAY05699	Mouse TNF receptor
27	397	19.0	76	22	AAM17149	Peptide #3583 enco
28	397	19.0	76	22	AAM29641	Peptide #3678 enco
29	397	19.0	76	22	AAM04844	Peptide #3526 enco
30	352	16.8	67	21	AAB26987	Human TNF receptor
31	130	6.2	2907	21	AAY57452	Human transcriptio
32	119	5.7	2781	21	AAY57453	Human transcriptio
33	115.5	5.5	2183	21	AAB03533	Murine factor V SE
34	111.5	5.3	355	22	AAM40196	Human polypeptide
35	111.5	5.3	521	21	AAB57089	Human prostate can
36	111.5	5.3	582	21	AAY44722	Human immune syste
37	110.5	5.3	290	21	AAB42240	Human ORFX ORF2004
38	105	5.0	1070	22	AAM00894	Human bone marrow
39	100	4.8	830	21	AAY96730	PRO539, a Costal-2
40	99	4.7	558	22	AAB85479	Human serine/threo
41	99	4.7	2343	20	AAW80989	Canine factor VIII
42	99	4.7	2343	21	AAY57846	Canine Factor VIII
43	98.5	4.7	700	12	AAR11354	Cellular Receptor
44	98	4.7	804	21	AAB42128	Human ORFX ORF1892
45	97	4.6	340	21	AAB15554	Apoptosis related

ALIGNMENTS

RESULT 1
AAY67947
ID AAY67947 standard; Protein; 405 AA.
XX
AC AAY67947;
XX
DT 27-MAR-2000 (first entry)
XX
DE Human DETH protein SEQ ID NO:2.
XX
KW Human; DETH; apoptosis; TNFR; neurodegenerative diseases; cancer;
KW death domain expressing tumour necrosis factor receptor family homologue;
KW autoimmune disease; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9966039-A1.
XX
PD 23-DEC-1999.
XX
PF 08-JUN-1999; 99WO-GB01793.
XX
PR 12-JUN-1998; 98GB-0012607.
XX
PA (ZENE) ZENECA LTD.
XX
PI Lu JJ, Gomes BC, Fieles WE;
XX
DR WPI; 2000-097744/08.
DR N-PSDB; AAZ57184.
XX
PT New protein having 80% identity to death domain expressing tumor
PT necrosis factor receptor family homologue -
XX

```
PS   Claim 5; Page 35-36; 40pp; English.
XX
CC   The present sequence represents a human death domain expressing tumour
CC   necrosis factor receptor family homologue, designated DETH. DETH has a
CC   role in apoptosis and appears to be member of the TNFR1/FAS/NGFR family
CC   of receptors. The DETH protein can be used for inducing apoptosis by
CC   expressing DETH in a cell. The protein is useful for identifying
CC   inhibitors and agonists of apoptosis that treat neurodegenerative
CC   diseases (inhibition), cancer and autoimmune diseases (induction).
CC   DETH-specific antibodies are useful for diagnosis of conditions and
CC   diseases associated with DETH expression. The protein avoids the use
CC   of harmful chemotherapy, where the loss of the p53 tumour repressor
CC   gene can lead to drug resistant tumour cells following treatment.
XX
SQ   Sequence   405 AA;


  Query Match           100.0%;  Score 2091;  DB 21;  Length 405;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-197;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 60

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 120

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 180

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 240

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 300

Qy     301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 360

Qy     361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     361 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 405


RESULT   2
AAY05678
ID   AAY05678 standard; Protein; 631 AA.
XX
AC   AAY05678;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Human full-length tumour necrosis factor receptor ZTNFR-6.
XX
KW   ZTNFR-6; tumour necrosis factor receptor-6; human;
KW   cell maturation; bone cell regulation.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..17
FT                   /note= "signal peptide"
FT   Protein         18..631
FT                   /note= "mature protein"
FT   Protein         18..188
FT                   /note= "a polypeptide comprising this region of the
FT                          sequence is claimed in Claim 28(b)"
FT   Protein         18..325
FT                   /note= "a polypeptide comprising this region of the
FT                          sequence is claimed in Claim 28(c)"
FT   Protein         37..188
FT                   /note= "a polypeptide comprising this region of the
FT                          sequence is claimed in Claim 28(a)"
FT   Protein         66..145
FT                   /note= "a polypeptide comprising this region of the
FT                          sequence is claimed in Claim 26(a)"
FT   Region          37..65
FT                   /note= "extracellular, cysteine-rich repeat"
FT   Region          66..107
FT                   /note= "extracellular, cysteine-rich repeat"
FT   Region          108..145
FT                   /note= "extracellular, cysteine rich repeat"
FT   Region          146..188
FT                   /note= "extracellular, cysteine-rich repeat"
FT   Region          189..325
FT                   /note= "linker region"
FT   Domain          326..346
FT                   /note= "transmembrane domain"
FT   Domain          347..631
FT                   /note= "cytoplasmic domain"
FT   Domain          390..471
FT                   /note= "death domain, specifically claimed in
FT                          Claim 32(a)"
XX
PN   WO9911790-A1.
XX
PD   11-MAR-1999.
XX
PF   03-SEP-1998;   98WO-US18364.
XX
PR   04-SEP-1997;   97US-0923725.
PR   04-SEP-1997;   97US-0057608.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Farrah TM,  Gross JA,  Matthews SM;
XX
DR   WPI; 1999-205190/17.
DR   N-PSDB; AAX25259.
XX
PT   New secreted or membrane bound tumor necrosis factor receptor
PT   ZTNFR-6  - useful for detecting a genetic abnormality in a patient
XX
PS   Claim 26; Page 113-115; 145pp; English.
XX
CC   The present sequence represents a novel human full-length
CC   membrane-bound tumour necrosis factor receptor, designated ZTNFR-6,
CC   that is characterised by 4 cysteine-rich pseudo-repeat motifs.
CC   ZTNFR-6 polypeptides can be obtained using recombinant techniques.
CC   A polynucleotide (see AAX25259) encoding ZTNFR-6 has been isolated
CC   from an EST database.  A secreted, soluble form (see AAY05679) of
CC   ZTNFR-6 has also been identified, which lacks the transmembrane
CC   and cytoplasmic domains of the membrane-bound protein.  ZTNFR-6
CC   polypeptides, including the isolated extracellular region,
CC   transmembrane domain and death domain, are claimed.  ZTNFR-6
CC   polypeptides are useful in methods that promote cellular maturation
CC   and bone cell regulation.  Antibodies raised against ZTNFR-6 are
CC   useful for immunohistochemical tagging of cells expressing ZTNFR-6
CC   for use in diagnosis, isolating ZTNFR-6, generating anti-idiotypic
CC   antibodies, and as neutralising antibodies or antagonists that block
CC   ZTNFR-6 in vivo and in vitro.  The polypeptides can also be used to
CC   identify agonists and antagonists of ZTNFR-6.  Agonists are useful
CC   for modifying the proliferation and development of target cells in
CC   vitro and in vivo.
XX
SQ   Sequence   631 AA;


  Query Match           100.0%;  Score 2091;  DB 20;  Length 631;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-196;
```

```
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    227 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 286

Qy     61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    287 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 346

Qy    121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    347 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 406

Qy    181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    407 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 466

Qy    241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    467 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 526

Qy    301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    527 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 586

Qy    361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
          |||||||||||||||||||||||||||||||||||||||||||||
Db    587 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 631


RESULT   3
AAW75792
ID   AAW75792 standard; Protein; 655 AA.
XX
AC   AAW75792;
XX
DT   21-DEC-1998  (first entry)
XX
DE   Human tumour necrosis factor related receptor TR7.
XX
KW   Tumour necrosis factor related receptor; TR7; human; inflammation;
KW   arthritis; septicaemia; autoimmune disease; psoriasis;
KW   inflammatory bowel disease; infection; graft-versus-host disease;
KW   transplant rejection; stroke; acute respiratory disease syndrome;
KW   ischaemia; restenosis; brain injury; AIDS; bone disease; cancer;
KW   atherosclerosis; Alzheimer's disease; diagnosis; therapy; vaccine.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "encoded by ACG"
XX
PN   EP869179-A1.
XX
PD   07-OCT-1998.
XX
PF   01-APR-1998;   98EP-0302528.
XX
PR   28-OCT-1997;   97US-0959382.
PR   02-APR-1997;   97US-0041796.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Deen KC,  Hurle MR,  Tan KB,  Young P;
XX
DR   WPI; 1998-508493/44.
DR   N-PSDB; AAV57441.
XX
PT   New tumour necrosis factor receptor TR7 polypeptides and
PT   polynucleotides - useful as diagnostic reagents and for treating
PT   Alzheimer's disease, AIDS and cancer
XX
PS   Claim 11; Page 19-21; 25pp; English.
XX
CC   This is the amino acid sequence of a novel member of the human
CC   tumour necrosis factor (TNF) receptor superfamily, termed TR7.
CC   It was deduced from the coding regions of overlapping isolated
CC   cDNA clones (see AAV57441).  Expression systems, host cells and a
CC   method of producing TR7 polypeptides are claimed.  TR7 polypeptides
CC   are useful for diagnosing diseases or susceptibility to diseases by
CC   determining TR7 polypeptide or mRNA expression.  TR7 polypeptide
CC   can be used to screen for agonists and antagonists which bind the
CC   receptor.  These can be used in treatment to inhibit or enhance TR7
CC   activity.  TR7 antibodies are generated using TR7 polypeptide
CC   fragments, and are used for treatment of diseases.  TR7 polypeptides
CC   and polynucleotides can be administered directly ás vaccines for
CC   prevention of diseases.  Diseases diagnosed, treated or prevented
CC   by the above methods include: chronic and acute inflammation,
CC   arthritis, septicemia, autoimmune diseases (e.g. inflammatory bowel
CC   disease, psoriasis), transplant rejection, graft vs. host disease,
CC   infection, stroke, ischaemia, acute respiratory disease syndrome,
CC   restenosis, brain injury, AIDS, bone diseases, cancer (e.g.
CC   lymphoproliferative disorders), atherosclerosis, and Alzheimer's
CC   disease.
XX
SQ   Sequence   655 AA;


  Query Match             100.0%;  Score 2091;  DB 19;  Length 655;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-196;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy     61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 370

Qy    121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430

Qy    181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy    241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy    301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy    361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
          |||||||||||||||||||||||||||||||||||||||||||||
Db    611 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 655


RESULT   4
AAY41693
ID   AAY41693 standard; Protein; 655 AA.
XX
AC   AAY41693;
XX
DT   07-DEC-1999  (first entry)
XX
DE   Human PRO868 protein sequence.
XX
KW   Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW   probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW   secreted protein; transmembrane protein.
```

XX
OS Homo sapiens.
XX
PN WO9946281-A2.
XX
PD 16-SEP-1999.
XX
PF 08-MAR-1999; 99WO-US05028.
XX
PR 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
PA (GETH) GENENTECH INC.
XX
PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
DR WPI; 1999-551358/46.
DR N-PSDB; AAZ33945.
XX
PT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders -
XX
PS Claim 12; Fig 26; 530pp; English.
XX
CC The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. AAZ33891 to
CC AAZ34338, and AAY41685 to AAY41774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
CC invention.
XX
SQ Sequence 655 AA;

Query Match 100.0%; Score 2091; DB 20; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy     61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 370

Qy    121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430

Qy    181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490
```

```
Qy    241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy    301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy    361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
          |||||||||||||||||||||||||||||||||||||||||||||
Db    611 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 655


RESULT   5
AAW81059
ID   AAW81059 standard; Protein; 655 AA.
XX
AC   AAW81059;
XX
DT   10-MAY-1999  (first entry)
XX
DE   Novel human tumor necrosis factor receptor TR9.
XX
KW   Human; tumour necrosis factor receptor; TNFR; TR9 receptor; cancer;
KW   apoptosis; agonist; inhibition; autoimmune disorder; viral infection;
KW   inflammation; antagonist; AIDS; neurodegenerative disorder.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..40
FT                   /note= "signal peptide"
FT   Peptide         41..655
FT                   /note= "mature protein"
XX
PN   WO9856892-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US11932.
XX
PR   11-JUN-1997;   97US-0052991.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Fan P,  Gentz RL,  Ni J,  Yu G;
XX
DR   WPI; 1999-060325/05.
DR   N-PSDB; AAV99927.
XX
PT   New isolated tumour necrosis factor-like receptor, TR9 - used to
PT   develop products for treating e.g. cancers, autoimmune disorders,
PT   viral infections, inflammation, graft rejection, neurodegenerative
PT   disorders or ischaemic injury
XX
PS   Claim 1; Fig 1; 134pp; English.
XX
CC   This is the amino acid sequence of the human tumour necrosis factor
CC   receptor (TNFR), TR9 receptor, used in the method of the invention
CC   to develop products to treat disorders such as cancers. The novel
CC   TNFR, TR9, can be used to identify agents for modifying apoptosis.
CC   Agonists can be used to treat diseases associated with increased
CC   cell survival, or the inhibition of apoptosis, including cancers
CC   (e.g. follicular lymphomas, carcinomas with p53 mutations, and
CC   hormone-dependent tumours, such as breast cancer, prostate cancer,
CC   Kaposi's sarcoma and ovarian cancer), autoimmune disorders (e.g.
CC   systemic lupus erythematosus and immune-related glomerulonephritis
CC   rheumatoid arthritis), viral infections (e.g. herpes viruses, pox
CC   viruses and adenoviruses), inflammation, graft vs host disease,
CC   acute graft rejection and chronic graft rejection. Antagonists
CC   can be used to treat diseases associated with increased apoptosis
CC   including AIDS, neurodegenerative disorders (e.g. Alzheimer's
CC   disease, Parkinson's disease, amyotrophic lateral sclerosis,
CC   retinitis pigmentosa, cerebellar degeneration), myelodysplstic
CC   syndromes (e.g. aplastic anemia), ischemic injury (e.g. that
CC   caused by myocardial infarction, stroke and reperfusion injury),
CC   toxin-induced liver disease (e.g. that caused by alcohol), septic
CC   shock, cachexia, anorexia, inflammatory diseases and stress response
CC   related diseases, such as inflammatory bowel disease, rheumatoid
CC   arthritis, osteoarthritis, psoriasis and septicemia. The products
CC   can also be used for detection and diagnosis.
XX
SQ   Sequence   655 AA;


  Query Match           100.0%;  Score 2091;  DB 20;  Length 655;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-196;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy     61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlflllvlvvivvcsi 370

Qy    121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430

Qy    181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy    241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy    301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy    361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
          |||||||||||||||||||||||||||||||||||||||||||||
Db    611 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 655


RESULT   6
AAB26981
ID   AAB26981 standard; Protein; 655 AA.
XX
AC   AAB26981;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Human tumour necrosis factor receptor TR9.
XX
KW   Human; tumour necrosis factor; TR9 receptor; immunosuppressive;
KW   antiinflammatory; cardiant; antiasthmatic; antidiabetic; antiallergic;
KW   antiathritic; antirheumatic; anti-HIV; anticonvulsant; cytostatic;
KW   neuroprotective; gene therapy; Death Domain Containing Receptor 6;
KW   common variable immunodeficiency; X-linked agammaglobulinaemia;
KW   severe combined immunodeficiency; Wiskott-Aldrich syndrome;
KW   autoimmune disease; rheumatoid arthritis; allergic encephalomyelitis;
KW   multiple sclerosis; diabetes mellitus; asthma; epilepsy; cancer;
KW   cardiovascular disease; neurological disease; protein coordinate data.
XX
OS   Homo sapiens.
XX
PN   WO200056862-A1.
XX
PD   28-SEP-2000.
XX
PF   16-MAR-2000; 2000WO-US06831.
XX
```

PR 24-MAR-1999; 99US-0126019.
PR 14-MAY-1999; 99US-0134220.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Gentz RL, Yu G, Fan P;
XX
DR WPI; 2000-594575/56.
DR N-PSDB; AAA99658.
XX
PT Nucleic acid molecule encoding a human tumor necrosis factor receptor,
PT known as TR9, useful for treating, preventing and diagnosing severe
PT combined immunodeficiency, autoimmune diseases, HIV infection, epilepsy
PT and cancer -
XX
PS Claim 20; Fig 1; 220pp; English.
XX
CC The present sequence is a novel human tumour necrosis factor
CC receptor, designated TR9. The TR9 receptor is also known as Death Domain
CC Containing Receptor 6. TR9 polypeptides, polynucleotides or agonists are
CC useful for treating, preventing or diagnosing common variable
CC immunodeficiency, X-linked agammaglobulinaemia, severe combined
CC immunodeficiency and Wiskott-Aldrich syndrome, autoimmune diseases (such
CC as rheumatoid arthritis, allergic encephalomyelitis, multiple sclerosis,
CC diabetes mellitus and asthma), HIV infection, epilepsy, cancer,
CC cardiovascular diseases and other neurological diseases.
XX
SQ Sequence 655 AA;

Query Match 100.0%; Score 2091; DB 21; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlflllvlvvivvcsi 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy     301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy     361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     611 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 655
```

RESULT 7
AAB01349
ID AAB01349 standard; Protein; 655 AA.
XX
AC AAB01349;
XX
DT 20-OCT-2000 (first entry)
XX
DE Tumour necrosis factor receptor homologue TRH1 (Clone 2733717).
XX
KW Tumour necrosis factor receptor homologue; TRH1; TNF; arthritis;
KW transplant rejection; activation; proliferation; differentiation;
KW apoptosis; immunosuppression; antiinflammatory; immunostimulation;
KW probe; primer; human.
XX
OS Homo sapiens.
XX
PN WO200034294-A2.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US29400.
XX
PR 11-DEC-1998; 98US-0111826.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Bowen MA, Siemers N;
XX
DR WPI; 2000-423364/36.
DR N-PSDB; AAA47395.
XX
PT Novel tumor necrosis factor receptor homologue-1 useful as a target for
PT immunosuppressive, antiinflammatory and/or immunostimulatory drug
PT development
XX
PS Claim 10; Fig 2a-2d; 42pp; English.
XX
CC The tumour necrosis factor receptor homologue TRH1 can be used for
CC treating a mammal e.g. a human, at risk for a disorder characterized
CC by an aberrant or unwanted level or biological activity of TRH1,
CC e.g. rheumatoid arthritis and transplant rejection. TRH1 may also be
CC useful to leach out or block a ligand which is found to bind to the
CC TRH1. TRH1 may be used in various drug screening techniques and to
CC identify fragments and analogs of a protein or peptide (agonist or
CC antagonist) which bind to TRH1. The TRH1 protein plays a role in
CC cellular function, cell activation, proliferation, differentiation,
CC and apoptosis. The interaction between the novel TNFr protein of the
CC present invention and intracellular signaling molecules and/or its
CC potential co-receptor may serve as a novel target for
CC immunosuppressive, antiinflammatory and/or immunostimulatory drug
CC development. Gene constructs can also be used as part of a gene
CC therapy protocol to deliver nucleic acids encoding the TRH1, or an
CC agonist or antagonist form of a TRH1 protein or peptide. Antibody
CC directed against TRH1 can be used to reject TRH1 in tissues
CC and cells. They can also be used to make targeted antibody that
CC destroy TRH1 expressing cells. Fragments of the TRH1 gene can be
CC used as diagnostic probes or as PCR primers. Fragments of the full
CC length gene may be used as hybridization probes for a cDNA library to
CC isolate the full length gene and to isolate other genes which have a
CC high sequence similarity. The probes may be used to identify a cDNA
CC clone corresponding to a full length transcript and a genomic clone
CC or clones that contain the complete gene including regulatory and
CC promoter regions, exons, and introns.
XX
SQ Sequence 655 AA;

Query Match 100.0%; Score 2091; DB 21; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlflllvlvvivvcsi 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430
```

```
Qy    181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy    241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy    301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy    361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
          |||||||||||||||||||||||||||||||||||||||||||||
Db    611 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 655
```

RESULT 8
AAB01338
ID AAB01338 standard; Protein; 655 AA.
XX
AC AAB01338;
XX
DT 25-SEP-2000 (first entry)
XX
DE DR-6 death receptor.
XX
KW UL144; death receptor; apoptosis; programmed cell death; FAS;
KW TNF-R1; TRAMP; DR-6; TRAIL; modulation; treatment; cancer; virus;
KW human.
XX
OS Homo sapiens.
XX
PN WO200034335-A2.
XX
PD 15-JUN-2000.
XX
PF 03-DEC-1999; 99WO-US26035.
XX
PR 04-DEC-1998; 98US-0205018.
XX
PA (SCHE) SCHERING CORP.
XX
PI Leong C, Phillips JH;
XX
DR WPI; 2000-423383/36.
XX
PT Purified or recombinant polypeptide for modulating apoptosis comprises
PT a sequence which binds to an antibody specific for UL144 or its
PT fragments
XX
PS Disclosure; Page 68-70; 76pp; English.
XX
CC A pure or recombinant polypeptide which binds to a polyclonal antibody
CC specific for the mature UL144 is useful for screening molecules which
CC block induction of apoptosis or interfere with antiapoptotic activity.
CC The polypeptide is also useful for modulating apoptosis and useful in
CC treatment of conditions associated with abnormal physiology or
CC development, such as cancer or degenerative conditions and for
CC regulation of viral infection and replication. At least five
CC different death receptors are known, which include the CD95
CC (Fas/APO-1), the TNF receptor-1, TNF receptor apoptosis-mediated
CC protein (TRAMP), death receptor-6 (DR-6), and TNF-related
CC apoptosis-inducing ligand (TRAIL) receptors 1, 2 and 4.
XX
SQ Sequence 655 AA;

Query Match 100.0%; Score 2091; DB 21; Length 655;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy     61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlflllvlvvivvcsi 370

Qy    121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430

Qy    181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy    241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy    301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy    361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
          |||||||||||||||||||||||||||||||||||||||||||||
Db    611 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 655
```

RESULT 9
AAY77460
ID AAY77460 standard; Protein; 655 AA.
XX
AC AAY77460;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human TNF receptor-like protein HSLJD37R, SEQ ID NO:10.
XX
KW TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
KW dermatological; antithyroid.
XX
OS Homo sapiens.
XX
PN WO200001817-A2.
XX
PD 13-JAN-2000.
XX
PF 06-JUL-1999; 99WO-US12366.
XX
PR 06-JUL-1998; 98US-0110938.
PR 13-JUL-1998; 98US-0114466.
PR 23-JUL-1998; 98US-0093897.
PR 12-AUG-1998; 98US-0132968.
PR 18-AUG-1998; 98US-0136214.
PR 11-SEP-1998; 98US-0099999.
XX
PA (SCHE) SCHERING CORP.
XX
PI Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
PI Bazan JF, Mahony D, Lees EM;
XX
DR WPI; 2000-171015/15.
DR N-PSDB; AAZ92406.
XX
PT New isolated mammalian genes, used to develop products for treating
PT e.g. immune, inflammatory or allergic abnormalities, cancers or
PT degenerative conditions -
XX
PS Claim 24; Page 163-165; 218pp; English.

```
XX
CC   The invention relates to a number of primate and/or rodent proteins, and
CC   the genes which encode them. The invention encompasses human dendritic
CC   cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC   factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
CC   human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub
CC   12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC   properties of ligands for proteins comprising a leucine-rich motif
CC   (LRR); human cyclin E2; cDNAs encoding these proteins; and antibodies
CC   against these proteins. The proteins can be used for modulating the
CC   physiology or development of a cell. They can be used to mediate uptake
CC   of substrates (e.g., prostaglandin-like molecules), to modulate or
CC   mediate cellular interactions (e.g., induce or prevent trafficking,
CC   proliferation, or differentiation of cells), or are intracellular
CC   proteins which are important in various cellular processes such as the
CC   deubiquitination of proteins or cell cycle regulation. The products can
CC   be used for treating medical conditions such as immune, inflammatory or
CC   allergic disorders, or abnormal cellular proliferation, for example,
CC   cancers or degenerative conditions. They can be used to modulate immune
CC   responses in disease states e.g., autoimmune disorders, including
CC   rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
CC   autoimmune thyroiditis, as well as acute and chronic inflammatory
CC   responses in which T cell activation, expansion, and/or immunological T
CC   cell memory play an important role. Sequences AAY77458-Y77461 and
CC   AAY77465-Y77468 represent TNF receptor family-related proteins. AAY77458
CC   is the human protein HDTEA84, AAY77459-Y77461 are human HSLJD37R
CC   proteins, AAY77465 is murine Rank-like protein RANKL, and AAY77466-Y77468
CC   are human RANKL proteins.
XX
SQ   Sequence   655 AA;

  Query Match           100.0%;  Score 2091;  DB 21;  Length 655;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-196;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy     301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy     361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     611 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 655


RESULT  10
AAY92846
ID   AAY92846 standard; Protein; 691 AA.
XX
AC   AAY92846;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Human osteoprotegrin-like 4 (OPGx4) protein.
XX
KW   OPGx4; osteoprotegrin-like; DR6 TNF-related death receptor; agonist;
KW   inhibitor; bone resorption; vascular calcification; apoptosis;
KW   osteopathic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..36
FT                   /note= "36 amino acid extension relative to
FT                          DR6 TNF-related death receptor"
XX
PN   WO200024771-A2.
XX
PD   04-MAY-2000.
XX
PF   22-OCT-1999;   99WO-US24913.
XX
PR   23-OCT-1998;   98US-0105481.
PR   01-OCT-1999;   99US-0156993.
PR   21-OCT-1999;   99US-0422680.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Yang M,  Lichenstein H,  McDonald WF;
XX
DR   WPI; 2000-350692/30.
DR   N-PSDB; AAA28728.
XX
PT   Nucleic acids encoding osteoprotegerin-like (OPGx) polypeptides, useful
PT   for treating disorders associated with bone metabolism, such as
PT   osteoporosis and osteopetrosis
XX
PS   Claim 24; Fig 2; 81pp; English.
XX
CC   Osteoprotegrin-like polypeptides (OPGx) are expressed in bone, lymph
CC   node, germinal B cells and kidney. There appear to be at least two splice
CC   variants of this gene, a transmembrane form and an extracellular domain
CC   form. Human OPGx1, 2 and 4 contain a 36 amino acid extension at their
CC   N-terminal relative to the DR6 tumour necrosis factor (TNF)-related death
CC   receptor. The OPGx polypeptides, agonists and antibodies are useful in
CC   methods to inhibit osteoclast-mediated bone resorption or vascular
CC   calcification and to modulate cell death (apoptosis). This is useful for
CC   treating disorders associated with bone metabolism, such as osteoporosis,
CC   osteopetrosis, or a condition characterized by loss of bone, breakdown of
CC   tissue, or excessive readsorption of bone tissue.
XX
SQ   Sequence   691 AA;

  Query Match           100.0%;  Score 2091;  DB 21;  Length 691;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-196;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     287 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 346

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     347 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 406

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     407 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 466

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     467 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 526

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    527 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 586

Qy    301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    587 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 646

Qy    361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
          |||||||||||||||||||||||||||||||||||||||||||||
Db    647 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 691


RESULT 11
AAB44249
ID   AAB44249 standard; Protein; 655 AA.
XX
AC   AAB44249;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human PRO868 (UNQ437) protein sequence SEQ ID NO:64.
XX
KW   Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;
KW   expressed sequence tag; detection; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200053756-A2.
XX
PD   14-SEP-2000.
XX
PF   18-FEB-2000; 2000WO-US04341.
XX
PR   08-MAR-1999;   99WO-US05028.
PR   12-MAR-1999;   99US-0123957.
PR   29-MAR-1999;   99US-0126773.
PR   21-APR-1999;   99US-0130232.
PR   28-APR-1999;   99US-0131445.
PR   14-MAY-1999;   99US-0134287.
PR   23-JUN-1999;   99US-0141037.
PR   26-JUL-1999;   99US-0145698.
PR   29-OCT-1999;   99US-0162506.
PR   30-NOV-1999;   99WO-US28313.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28565.
PR   16-DEC-1999;   99WO-US30095.
PR   30-DEC-1999;   99WO-US31243.
PR   30-DEC-1999;   99WO-US31274.
PR   05-JAN-2000; 2000WO-US00219.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Baker KP,  Botstein D,  Desnoyers L,  Eaton DL;
PI   Ferrara N,  Filvaroff E,  Fong S,  Gao W,  Gerber H,  Gerritsen ME;
PI   Goddard A,  Godowski PJ,  Grimaldi CJ,  Gurney AL,  Hillan KJ;
PI   Kljavin IJ,  Kuo SS,  Napier MA,  Pan J,  Paoni NF,  Roy MA;
PI   Shelton DL,  Stewart TA,  Tumas D,  Williams PM,  Wood WI;
XX
DR   WPI; 2000-611443/58.
DR   N-PSDB; AAC78474.
XX
PT   Novel PRO polypeptides and polynucleotides used in detection methods,
PT   to target bioactive molecules to specific cells, and to modulate
PT   cellular activities -
XX
PS   Claim 12; Fig 26; 636pp; English.
XX
CC   AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC   sequence tag) sequences which encode secreted or transmembrane PRO
CC   polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC   activity. The polynucleotides and polypeptides can be used for detecting
CC   the presence of PRO polypeptides in samples, for linking bioactive
CC   molecules to cells and for modulating biological activities of cells,
CC   using the polypeptides for specific targeting. The polypeptide targeting
CC   can be used to kill the target cells, e.g. for the treatment of cancers.
CC   The polypeptide pairs provide specific targeting of bioactive molecules
CC   to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC   the isolation of the PRO polynucleotide sequences.
XX
SQ   Sequence   655 AA;


  Query Match             99.9%;  Score 2089;  DB 21;  Length 655;
  Best Local Similarity   99.8%;  Pred. No. 1.9e-196;
  Matches  404;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy     61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 370

Qy    121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db    371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilkivaaqvgsqw 430

Qy    181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy    241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy    301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy    361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
          |||||||||||||||||||||||||||||||||||||||||||||
Db    611 elrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 655


RESULT 12
AAB93023
ID   AAB93023 standard; Protein; 655 AA.
XX
AC   AAB93023;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:11787.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
```

```
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 11787; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   655 AA;

  Query Match             99.8%;  Score 2087;  DB 22;  Length 655;
  Best Local Similarity   99.8%;  Pred. No. 3e-196;
  Matches  404;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqdknkgff 550

Qy     301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 vdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhflnpe 610

Qy     361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||| ||||||||||||
Db     611 elrvieeipqaedkldrlfeiigvksqeasqtfldsvyshlpdll 655


RESULT  13
AAY05695
ID   AAY05695 standard; Protein; 573 AA.
XX
AC   AAY05695;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Mouse TNF receptor TRL.
XX
KW   TRL; TNF receptor-like; tumour necrosis factor receptor; mouse;
KW   signal transduction; cell differentiation; prostate cancer;
KW   inflammation; arthritis; diabetes; insulin resistance; diagnosis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Domain          8..49
FT                   /note= "cysteine-rich domain"
FT   Domain          50..87
FT                   /note= "cysteine-rich domain"
FT   Domain          88..130
FT                   /note= "cysteine-rich domain"
FT   Domain          270..288
FT                   /note= "transmembrane domain"
FT   Domain          333..416
FT                   /note= "death domain"
XX
PN   WO9915663-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US20219.
XX
PR   17-MAR-1998;   98US-0042785.
PR   26-SEP-1997;   97US-0938896.
XX
PA   (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI   Busfield SJ;
XX
DR   WPI; 1999-254712/21.
DR   N-PSDB; AAX25320.
XX
PT   Nucleic acid encoding tumor necrosis factor receptor-like protein
XX
PS   Claim 1d; Page 113-116; 170pp; English.
XX
CC   The present sequence represents murine TRL (TNF receptor-like), a
CC   novel member of the tumour necrosis factor receptor superfamily.
CC   The sequence is predicted from the sequence of a cDNA clone (see
CC   AAX25320) from a murine megakaryocyte cDNA library.  Murine TRL is
CC   a 63 kDa membrane-bound protein that includes 3 cysteine-rich
CC   domains and a C-terminal death domain.  Human TRLs (see AAY05696-97)
CC   have also been identified.  The invention provides full-length TRL
CC   proteins, TRL fusion proteins, antigenic peptides and anti-TRL
CC   antibodies, as well as TRL nucleic acids, recombinant expression
CC   vectors, host cells and non-human transgenic animals.  TRLs are
CC   regulators or modulators of cellular signal transduction, cellular
CC   proliferation or differentiation, cell survival and apoptosis,
CC   immune system cells, and cells involved in insulin resistance or
CC   the diabetic response.  They can be used to isolate cognate
CC   ligands, to modulate ligand/TRL interactions, to screen for
CC   potential modulators, and to treat conditions associated with
CC   abnormal TRL levels.  Antibodies, TRL protein mutants etc. that
CC   modulate activity of TRL are also used to treat/prevent conditions
CC   associated with aberrant TRL protein or expression, e.g. cell
CC   proliferation and differentiation disorders (particularly cancer,
CC   especially of the prostate), inflammation (e.g. arthritis) and
CC   metabolic disease (e.g. diabetes and insulin resistance).
XX
SQ   Sequence   573 AA;

  Query Match             91.5%;  Score 1914;  DB 20;  Length 573;
  Best Local Similarity   92.9%;  Pred. No. 2.6e-179;
```

```
  Matches  377;  Conservative    6;  Mismatches   21;  Indels    2;  Gaps    2;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||:|||:|||| || | |:||||||||||  |||  |:||||   | ||| |||||||
Db     169 mnstdsnstasvrtkvpsgieegtvpdntssts gkegtnrtlpnppqvthqqaphhrhil 228

Qy      61 KLLP-SMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCS 119
           |||| ||||| |||||| || |||||||||  ||||||||||||||||||||||||:|||||
Db     229 kllpssmeat-geksstaikapkrghprqnahkhfdinehlpwmivlfllvlvlivvcs 287

Qy     120 IRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQ 179
           ||||||||||||||||||||||||||||:|||||||||||| ||||||||||||||||||
Db     288 irkssrtlkkgprqdpsaivekaglkksltptqnrekwiyyrnghgidilklvaaqvgsq 347

Qy     180 WKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND 239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     348 wkdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrnd 407

Qy     240 VVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGF 299
           |||||||||||||||||||||||||||||||||||| ||||| |||||||| |||| |
Db     408 vvekirglmedttqletdklalpmspsplspspipspnvklenstlltvepspldknkcf 467

Qy     300 FVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNP 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     468 fvdesepllrcdstssgssalsrngsfitkekkdtvlrqvrldpcdlqpifddmlhilnp 527

Qy     360 EELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     528 eelrvieeipqaedkldrlfeiigvksqeasqtlldsvyshlpdll 573


RESULT  14
AAY05697
ID   AAY05697 standard; Protein; 605 AA.
XX
AC   AAY05697;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Human TNF receptor TRL II.
XX
KW   TRL II; TNF receptor-like; tumour necrosis factor receptor; human;
KW   signal transduction; cell differentiation; prostate cancer;
KW   inflammation; arthritis; diabetes; insulin resistance; diagnosis;
KW   therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..41
FT                   /note= "signal peptide" .
FT   Protein         42..605
FT                   /note= "mature protein"
FT   Domain          47..89
FT                   /note= "cysteine-rich domain"
FT   Domain          90..131
FT                   /note= "cysteine-rich domain"
FT   Domain          132..169
FT                   /note= "cysteine-rich domain"
FT   Domain          170..212
FT                   /note= "cysteine-rich domain"
FT   Domain          352..370
FT                   /note= "transmembrane domain"
FT   Domain          415..498
FT                   /note= "death domain"
XX
PN   WO9915663-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US20219.
XX
PR   17-MAR-1998;   98US-0042785.
PR   26-SEP-1997;   97US-0938896.
XX
PA   (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI   Busfield SJ;
XX
DR   WPI; 1999-254712/21.
DR   N-PSDB; AAX25322.
XX
PT   Nucleic acid encoding tumor necrosis factor receptor-like protein
XX
PS   Claim 1d; Page 154-157; 170pp; English.
XX
CC   The present sequence represents human TRL II (TNF receptor-like), a
CC   novel member of the tumour necrosis factor receptor superfamily.
CC   The sequence is predicted from the sequence of a cDNA clone (see
CC   AAX25322) isolated from an HeLa cell cDNA library.  TRL II is a
CC   a 66.2 kDa membrane-bound protein includes 4 cysteine-rich domains
CC   and a C-terminal unique region containing a death domain.  An
CC   alternatively spliced secreted form, human TRL I (see AAY05696),
CC   has also been identified.  The invention provides full-length TRL
CC   proteins, TRL fusion proteins, antigenic peptides and anti-TRL
CC   antibodies, as well as TRL nucleic acids, recombinant expression
CC   vectors, host cells and non-human transgenic animals.  TRLs are
CC   regulators or modulators of cellular signal transduction, cellular
CC   proliferation or differentiation, cell survival and apoptosis,
CC   immune system cells, and cells involved in insulin resistance or
CC   the diabetic response.  They can be used to isolate cognate
CC   ligands, to modulate ligand/TRL interactions, to screen for
CC   potential modulators, and to treat conditions associated with
CC   abnormal TRL levels.  Antibodies, TRL protein mutants etc. that
CC   modulate activity of TRL are also used to treat/prevent conditions
CC   associated with aberrant TRL protein or expression, e.g. cell
CC   proliferation and differentiation disorders (particularly cancer,
CC   especially of the prostate), inflammation (e.g. arthritis) and
CC   metabolic disease (e.g. diabetes and insulin resistance).
XX
SQ   Sequence   605 AA;


  Query Match             73.0%;  Score 1527;  DB 20;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-141;
  Matches  294;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 mnstesnssasvrpkvlssiqegtvpdntssargkedvnktlpnlqvvnhqqgphhrhil 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 kllpsmeatggeksstpikgpkrghprqnlhkhfdinehlpwmivlfllvlvvivvcsi 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 rkssrtlkkgprqdpsaivekaglkksmtptqnrekwiyycnghgidilklvaaqvgsqw 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQD 294
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 vekirglmedttqletdklalpmspsplspspipspnaklensalltvepspqd 544


RESULT  15
AAY92845
ID   AAY92845 standard; Protein; 600 AA.
XX
AC   AAY92845;
XX
DT   29-AUG-2000  (first entry)
XX
```

DE Murine osteoprotegrin like 5 (OPGx5) protein.
XX
KW OPGx5; osteoprotegrin-like; DR6 TNF-related death receptor; agonist;
KW inhibitor; bone resorption; vascular calcification; apoptosis;
KW osteopathic.
XX
OS Mus sp.
XX
PN WO200024771-A2.
XX
PD 04-MAY-2000.
XX
PF 22-OCT-1999; 99WO-US24913.
XX
PR 23-OCT-1998; 98US-0105481.
PR 01-OCT-1999; 99US-0156993.
PR 21-OCT-1999; 99US-0422680.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Yang M, Lichenstein H, McDonald WF;
XX
DR WPI; 2000-350692/30.
DR N-PSDB; AAA28727.
XX
PT Nucleic acids encoding osteoprotegerin-like (OPGx) polypeptides, useful
PT for treating disorders associated with bone metabolism, such as
PT osteoporosis and osteopetrosis
XX
PS Claim 23; Fig 9B; 81pp; English.
XX
CC Osteoprotegrin-like polypeptides (OPGx) are expressed in bone, lymph
CC node, germinal B cells and kidney. There appear to be at least two splice
CC variants of this gene, a transmembrane form and an extracellular domain
CC form. Human OPGx1, 2 and 4 contain a 36 amino acid extension at their
CC N-terminal relative to the DR6 tumour necrosis factor (TNF)-related death
CC receptor. The OPGx polypeptides, agonists and antibodies are useful in
CC methods to inhibit osteoclast-mediated bone resorption or vascular
CC calcification and to modulate cell death (apoptosis). This is useful for
CC treating disorders associated with bone metabolism, such as osteoporosis,
CC osteopetrosis, or a condition characterized by loss of bone, breakdown of
CC tissue, or excessive readsorption of bone tissue.
XX
SQ Sequence 600 AA;

Query Match 66.8%; Score 1396; DB 21; Length 600;
Best Local Similarity 89.4%; Pred. No. 2.6e-128;
Matches 271; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

```
Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||:|||:|||| || | |:|||||||||    |||  |:||||   | ||| |||||||
Db     251 mnstdsnstasvrtkvpsgieegtvpdntsptsgkegtnrtlpnppqvthqqaphhrhil 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           |||||     |||||| || |||||||||-||||||||||||||||||||||:||||||
Db     311 kllpssmkatgeksstaikapkrghprqnahkhfdinehlpwmivlflllvlvlivvcsi 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||:|||||||||||| ||||||||||||||||||
Db     371 rkssrtlkkgprqdpsaivekaglkksltptqnrekwiyyrnghgidilklvaaqvgsqw 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 kdiyqflcnaserevaafsngytadherayaalqhwtirgpeaslaqlisalrqhrrndv 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           |||||||||||||||||||||||||||||||||||| ||||| |||||||| |||| ||
Db     491 vekirglmedttqletdklalpmspsplspspipspnvklenstlltvepspldknkcff 550

Qy     301 VDE 303
           |||
Db     551 vde 553
```

Search completed: January 4, 2002, 08:37:53
Job time: 102 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:37 ; Search time 15.98 Seconds
(without alignments)
1930.582 Million cell updates/sec

Title: US-09-733-956-2
Perfect score: 2091
Sequence: 1 MNSTESNSSASVRPKVLSSI.....SQEASQTLLDSVYSHLPDLL 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	115.5	5.5	2183	2	T42764	coagulation factor
2	115	5.5	927	2	T51536	hypothetical prote
3	114.5	5.5	290	2	T08692	hypothetical prote
4	112	5.4	1822	2	T14106	probable GTPase-ac
5	107	5.1	309	2	T17557	procyclin homolog
6	107	5.1	461	2	T22946	hypothetical prote
7	105.5	5.0	452	2	T46147	zinc finger protei
8	105.5	5.0	524	2	E71881	hypothetical prote
9	105.5	5.0	1405	2	T04426	hypothetical prote
10	104.5	5.0	607	2	T40906	probable ATP-depen
11	104.5	5.0	3140	2	S47508	genome polyprotein
12	103.5	4.9	1301	2	S51323	SAC3 protein - yea
13	103	4.9	1231	2	T18532	serine/threoine pr
14	103	4.9	2176	2	T13806	toucan gene protei
15	101.5	4.9	494	2	T28067	hypothetical prote
16	101.5	4.9	514	2	T15338	hypothetical prote
17	101.5	4.9	527	2	B64633	hypothetical prote
18	101	4.8	342	2	E82955	TonB protein PA553
19	100	4.8	498	2	F83523	probable colicin-l
20	100	4.8	1087	2	T49496	hypothetical prote
21	99.5	4.8	872	2	T50369	probable serine/th
22	99	4.7	604	2	T41249	DEAD box ATP-depen
23	98.5	4.7	3140	1	GNVSRA	genome polyprotein
24	98.5	4.7	26926	1	I38344	titin, cardiac mus
25	98	4.7	670	2	T13739	probable hormone r
26	98	4.7	897	2	S61137	probable membrane
27	98	4.7	1188	2	T20333	hypothetical prote
28	97	4.6	764	2	T05409	hypothetical prote
29	96.5	4.6	1643	2	T07961	myosin heavy chain
30	96.5	4.6	1940	2	T42715	ankyrin 3, splice
31	96.5	4.6	1943	2	T42713	ankyrin 3, splice
32	96.5	4.6	1961	2	T42716	ankyrin 3, splice
33	96	4.6	607	2	S27776	80K protein (allel
34	96	4.6	856	2	G71133	probable alpha-man
35	96	4.6	969	2	T41707	probable pheromone
36	95.5	4.6	733	2	S61984	probable membrane
37	95.5	4.6	962	2	A84497	probable retroelem
38	95.5	4.6	1025	2	S54044	probable membrane
39	95.5	4.6	1086	2	S16752	major merozoite su
40	95.5	4.6	1502	1	RGBYH1	CYC1/CYP3 transcri
41	95	4.5	707	2	S57157	hypothetical prote
42	95	4.5	3141	1	GNVSPD	genome polyprotein
43	94.5	4.5	1033	2	S19247	cell adhesion prot
44	94	4.5	249	2	T07300	hypothetical prote
45	94	4.5	294	2	G84504	probable VSP-1-lik

ALIGNMENTS

RESULT 1
T42764
coagulation factor V - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C;Accession: T42764
R;Yang, T.L.; Cui, J.; Rehumtulla, A.; Yang, A.; Moussalli, M.; Kaufman, R.J.; Ginsbu
Blood 91, 4593-4599, 1998
A;Title: The structure and function of murine factor V and its inactivation by protei
A;Reference number: Z22270; MUID:98282202
A;Accession: T42764
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2183 <YAN>
A;Cross-references: EMBL:U52925; NID:g3219690; PID:g3219691; PIDN:AAC99553.1
C;Function:
A;Pathway: blood coagulation
C;Superfamily: coagulation factor V; discoïdin I amino-terminal homology; ferroxidase
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasm
F;350-682/Domain: ferroxidase repeat homology <FOX1>
F;1541-1864/Domain: ferroxidase repeat homology <FOX2>

Query Match 5.5%; Score 115.5; DB 2; Length 2183;
Best Local Similarity 19.9%; Pred. No. 3;
Matches 97; Conservative 49; Mismatches 197; Indels 145; Gaps 20;

```
Qy       5 ESNSSASVRPKVLSSIQEGTVPD--NTSSARGKEDVNKTLPNL----------QVVNHQQ 52
           :||||     ::|| |    : |   |    |       | ||             |:
Db     773 DSNSS-----RILSKIINNNLKDFQRTLPGSGATVAGTLLRNLIGLDENFVLNSSTEHRS 827

Qy      53 GPHHRH----------ILKLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHF-------D 95
             :| :          :: |||      | :     |  | | |    |:
Db     828 SSYHENDMENPQSNITMVYLLPLGPKGSGNREQDKPKTIKTGRPHMMKHRFSWMKAPAGK 887

Qy      96 INEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNRE 155
             |                  |  |:| :  |     || ::    ||: :|
Db     888 TGRH-----------------SNPKNSYSGMKSEEDIPSELIP---LKQKITSKFLNR 925

Qy     156 KW----------IYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTAD 205
           :|          |   ||   |: ||                |: : :      | :
Db     926 RWRVASEKGSYEIIAANGEDTDVDKLT---------------NSPQNQNITVPRGESTS 969

Qy     206 HERAYAALQHWTIRGPE-----ASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDKLA 260
           |        | | |     : :     :||   |   :| :  :    : :  |||
Db     970 HTN--------TTRKPSDLPTFSGVGHKSPHVRQEEENSGFQKRQLFIRTRKKKKNKKLA 1021

Qy     261 L--PMSPSPLSP----SPIPSPNAKLENSALL-------TVEPSPQDKNKGFFVDESEP- 306
           | |:||    |    :  | |: :| | :||        : |     :     | | |
Db    1022 LHSPLSPRGFDPLRGHNHSPFPDRRLLNHSLLLHKSNETALSPDLNQTSPSMSTDRSLPD 1081

Qy     307 --LLRCDSTSSGSSALSRNGSFITKEKKDTVLRQ-----------VRLDPCDLQPIFD- 351
```

```
              : |   ||:|    |   :|   |   |             |  | :|    |
Db   1082 YNQYSKNDTEQMSSSLDLYQSVPAEEHSPTFPAQDPDQTHSTTDPSYRSSPPELSQGLDY 1141

Qy    352 DMLHFLNPEELRVIEEIPQAEDKLD--------------RLFEIIGVKSQEASQTLLDSV 397
          |: |   |::: :    |    |                 || |    | |  ||:   :
Db   1142 DLSHDFYPDDIGLTSFFPDQSQKSSFSSDDDQAIPSSDLSLFTI----SPELDQTI---I 1194

Qy    398 YSHLPDLL 405
          |  |  ||
Db   1195 YPDLDQLL 1202


RESULT   2
T51536
hypothetical protein T20K14_190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C;Accession: T51536
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-927 <SAT>
A;Cross-references: EMBL:AL391143
A;Experimental source: cultivar Columbia; BAC clone T20K14
C;Genetics:
A;Map position: 5
A;Introns: 51/1; 69/3; 620/3; 668/1
A;Note: T20K14_190


  Query Match             5.5%;  Score 115;  DB 2;  Length 927;
  Best Local Similarity   18.9%;  Pred. No. 0.99;
  Matches   83;  Conservative   51;  Mismatches  158;  Indels  146;  Gaps   15;

Qy     17 LSSIQEGTVPDN---TSSARGKEDVNKTLPNLQVVNHQQGPHHRHILKL---------- 62
          | |:  |   ||   |: :  |::  |:         |:|       :|
Db     45 LKSLPSGKASDNVGDTNISADKKETEKSKKKKTAKEKQRGVSSESSSRLSFSSSPCSSSF 104

Qy     63 --------LPSMEATGGEKSSTPIKGPKRGHPR-QNLHKHFDINEHLPWMIVLFLLLVLV 113
                  |  |      |:: |  | ||   |    || |
Db    105 SSADISTTASQFEQPGLSNGENPVREPTNGSPRWGGLMMPSDIRE--------------- 149

Qy    114 VIVVCSIRKSSRTL-KKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLV 172
           :|  || | :||  ::   | | :      | |  :|::|  :|       |  ::||
Db    150 -LVRSSIHKETRTRDEEALSQQPKSARANVSLLKESSPSRNSNEW-----SEGRRVVKL- 202

Qy    173 AAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISAL 232
                ||  :|  :  |||
Db    203 --------KDSPRF--SYDERE------------------------------------ 214

Qy    233 RQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIP-----------SPNAKL 280
                  |    :::| :|  |  :     :  | || |             |  |||
Db    215 --------TRKTGAKLKETPRLSLDSRSNSFRSARSSCSPEPQELVTGHRRTTSSVVAKL 266

Qy    281 ENSALLTVEPSPQDKNKGFFVDESEPLLRCDSTSSGSSALSRNGSFITKEKK-------- 332
             ::  ||      :  | |   |      ||     | |:  | : :|
Db    267 MGLEVIPDEPVTIQNRENRFCDSPRP------TSRVEVDLQRSRGFDSIKKMMPAKFPMK 320

Qy    333 -------DTVLRQVRLDPCDLQPIFDDM------LHFLNPE-ELRVIEEIPQAEDKLDRL 378
                |    ||::       :: ::      | |   | :|| :::| :| :|  :|
Db    321 ASPWAQVDGAKNQVKIPDATTLTVYGEIQKRLSQLEFKKSEKDLRALKQILEAMEKTQQL 380

Qy    379 FEIIGVKSQEASQTLLDS 396
                |  : ::||  |
Db    381 IS----KDDDDNKTLCSS 394


RESULT   3
T08692
hypothetical protein DKFZp564K112.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
C;Accession: T08692
R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16470
A;Accession: T08692
A;Molecule type: mRNA
A;Residues: 1-290 <DUE>
A;Cross-references: EMBL:AL049996
A;Experimental source: fetal brain; clone DKFZp564K112
C;Genetics:
A;Note: DKFZp564K112.1


  Query Match             5.5%;  Score 114.5;  DB 2;  Length 290;
  Best Local Similarity   21.1%;  Pred. No. 0.22;
  Matches   66;  Conservative   46;  Mismatches  104;  Indels   97;  Gaps   13;

Qy      2 NSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHILK 61
            :|:|   | : | ||   ||:  : : : :::   |:       ::   ::
Db      6 SSSEDAESLAPRSKVKKIIQKDIIKEASEASKENRDI-------EISTEEE--------- 49

Qy     62 LLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLP--------WMIVLFLLLVLV 113
               : ||  | |: :|  || | :::  :|    | :|         || :    | |
Db     50 -----KDTGDLKDSSLLK-TKRKHKKKHKERHKMGEEVIPLRVLSKSEWMDLKKEYLALQ 103

Qy    114 VIVVCSIRKSSRTLK-KGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYC------NGHGI 166
             : |::|:   :| :   :  | : :  |:|     | |||:     |      |  |
Db    104 KASMASLKKTISQIKSESEMETDSGVPQNTGMKNE--KTANREE----CRTQEKVNATGP 157

Qy    167 D-----ILKLVAAQVGSQWKDIYQFLCNASE--------------------REVAAFSNG 201
                |:|::: :     | :   |   ||                       :  |  |
Db    158 QFVSGVIVKIISTEPLPGRKQVRDTLAAISEVLYVDLLEGDTECHARFKTPEDAQAVINA 217

Qy    202 YT--------------ADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRGL 247
          ||               |||: |     |          | |   || : |   || ||
Db    218 YTEINKRHCWKLEILSGDHEQRY-----W----------QKILVDRQAKLNQPREKKRGT 262

Qy    248 MEDTTQLETDKLA 260
            :  || |  :||
Db    263 EKLITQAEKIRLA 275


RESULT   4
T14106
probable GTPase-activating protein SPA-1 - rat
N;Alternate names: protein p1294
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14106
R;Takeuchi, M.; Ide, N.; Hata, Y.; Takai, Y.
submitted to the EMBL Data Library, September 1997
A;Description: SPA-1 like protein identified through yeast two-hybrid screening using
A;Reference number: Z17877
A;Accession: T14106
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1822 <TAK>
A;Cross-references: EMBL:AF026504; NID:g2555182; PID:g2555183; PIDN:AAB81526.1


  Query Match             5.4%;  Score 112;  DB 2;  Length 1822;
  Best Local Similarity   20.0%;  Pred. No. 4.3;
  Matches   91;  Conservative   65;  Mismatches  152;  Indels  148;  Gaps   19;

Qy      2 NSTESNSSASVRPKVLSSIQEGTVPDNTSSARG--KEDVNKTLPNLQVVNHQQGPHHRHI 59
          ||| | : |:      :|| :    |   ||||   ||::: |  :
Db   1429 NSTFSINDATSHTSTMSS-RHSASPVVFSSARSSPKEELHPTTSS--------------- 1472

Qy     60 LKLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHF------------------------ 94
            :| ||  :     ||:   ||:  :|||     :
Db   1473 -QLAPSFSS-----SSSSSSGPRTFYPRQGATSKYLIGWKKPEGTINSVGFMDTRKRHQS 1526
```

```
Qy      95 DINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSM---TPT 151
           | ||       |    :| |:| |:  |:    : :|:  ||| :   :||
Db    1527 DGNE---------------IAHTRLRASTRDLRASPKPTSKSTIEE-DLKKLIDLESPT 1569

Qy     152 QNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCN----ASEREVAAFSNGYTADH- 206
             :|   :   ||:   :       :  : ::: | :    :|:||    |     |
Db    1570 PESQKNFKF---HGLSSPQSPFPSTPTSRRALHRTLSDESIYSSQREHFFTSRASLLDQA 1626

Qy     207 -------ERAYAAL-QHWTIRGPEASLA----------QLISALRQHRRNDVVEKIRG 246
                 | :| :   :| |  :|              :: ::: ||  : :   |
Db    1627 LPNDVFFSSTYPSLPKSLPLRRPSYTLGMKSLHGEFFASDSSLTDIQETRRQPIPDP--G 1684

Qy     247 LMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESE- 305
           ||                    |:|   : |: | |:    : : :   ||    |
Db    1685 LM---------------------PLPDTASDLDWSNLVDAAKAYEVQRASFFAASDEN 1721

Qy     306 --PLLRCDSTSSGSSAL----------SRNGSFITKEKKDTVLRQVRLDPCDLQPIF 350
             ||    | :| |  |            |:: |     | | :   :::   ||:
Db    1722 HRPL----SAASNSDQLEEQALVQMKSYSSSKDSSPTLASKVDQLEGMLKMLREDLKKEK 1777

Qy     351 DDMLHF------LNPEELRVIEEIPQAEDKLDRLFE 380
           :|  |       |  : ||: ||   | ||| :  |
Db    1778 EDKAHLQAEVEHLREDNLRLQEESQNASDKLKKFTE 1813


RESULT   5
T17557
procyclin homolog A67R - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17557
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-309 <GRA>
A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96435.1
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Gene: A67R


  Query Match             5.1%;  Score 107;  DB 2;  Length 309;
  Best Local Similarity   25.6%;  Pred. No. 0.89;
  Matches   52;  Conservative   21;  Mismatches   68;  Indels   62;  Gaps    9;

Qy     106 LFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHG 165
           || |  |  :       |||||                  |  |||       |:   ||
Db       7 LFALFALFTVA------SSRTL-----------------KYTTPT------IHNNCGHS 36

Qy     166 IDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRG----- 220
           :  :||: ||:  :  |   | |: |: |    : |          ::  |  ||
Db      37 VTGIKLINAQLTGE--DAPDFECDYSDNEQIVLTPGPN--------SVDKWA-RGVSIVA 85

Qy     221 -PEASLAQLI------,--SALRQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSP 270
            ||     ::       |:| :    |      |:|  |  : :|      | |   |
Db      86 IPELKSGDILNVNDPNPSGSSLGESCYTD------GVMMQTFDIASDNDVYLCSMSSSEP 139

Qy     271 SPIPSPNAKLENSALLTVEPSPQ 293
           || |||    | |   : ||||:
Db     140 SPEPSPEPSPEPSPEPSPEPSPE 162


RESULT   6
T22946
hypothetical protein F58G6.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T22946
R;Lloyd, C.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19641
A;Accession: T22946
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-461 <WIL>
A;Cross-references: EMBL:Z68217; PIDN:CAA92465.1; GSPDB:GN00022; CESP:F58G6.1
A;Experimental source: clone F58G6
C;Genetics:
A;Gene: CESP:F58G6.1
A;Map position: 4
A;Introns: 19/3; 57/3; 96/3; 130/3; 157/3; 186/3; 219/3; 271/3; 302/3; 335/3; 357/3
C;Superfamily: amphiphysin; RVS161 protein homology


  Query Match             5.1%;  Score 107;  DB 2;  Length 461;
  Best Local Similarity   20.8%;  Pred. No. 1.6;
  Matches   60;  Conservative   46;  Mismatches  100;  Indels   82;  Gaps   12;

Qy     119 SIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGS 178
           | : |  ::|   ::|    : || ::      |  |:     |   ::|| |     :
Db     143 SAKNSFNSVKASSKKDNDPKLAKATME-----LQAAEQMYTEMNNELLEILPAVFDSRIT 197

Qy     179 QWKDIYQFLCNAS---EREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQH 235
             : |  | | ||:   : : : |        ::   :: :  :  |||     |
Db     198 FFVDTLQTLFNANSVYQTDASKFHKQIVMQLDKLGESMDYLRVARPEARALTPI------ 251

Qy     236 RRNDVVEKIRGLMEDTTQL----------ETDKLALPMSPSP-LSPSPIPSPNAKLEN 282
                      ||| |            ::|  :|  :||| : |||:|:| ||
Db     252 -------------DTTSLASSDAPATPDQNKSDNSSLRQTPSPSVLPSPVPAPMAKPRE 297

Qy     283 SALLTV--------------EPSPQDKNKGFFVDESE--PLLRCDSTSSGSSALSRNGS 325
           | :                | :| |  : |   |::  | |    |::   |::
Db     298 SVVSMANSTNPFDDDDAEHDKEATPTDVEEKF---EAKVYPKLNAQQTAAAEQAIAA--- 351

Qy     326 FITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQAED 373
              :|||:        :|      |||       :|   :|| |: :|
Db     352 -AKREKKEA------SNP------FDD-----EDDESTEVEEAPKKKD 381


RESULT   7
T46147
zinc finger protein - Arabidopsis thaliana
N;Alternate names: protein T3A5.80
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46147
R;Bloecker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetier, F.; S
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23024
A;Accession: T46147
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-452 <BLO>
A;Cross-references: EMBL:AL132979
A;Experimental source: cultivar Columbia; BAC clone T3A5
C;Genetics:
A;Map position: 3
A;Introns: 44/1; 176/2; 243/1
A;Note: T3A5.80


  Query Match             5.0%;  Score 105.5;  DB 2;  Length 452;
  Best Local Similarity   21.4%;  Pred. No. 2;
  Matches   83;  Conservative   51;  Mismatches  114;  Indels  139;  Gaps   22;

Qy       2 NSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHILK 61
           ||:  :  |||   : |:  :  :|::|    ||:   : || :             ::
Db       7 NSSTVSGEASV--SISSTGNQNPLPNST----GKK--KRNLPGMP-------DPESEVIA 51

Qy      62 LLP-SMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           | | :: ||          || :|    |||  |     :||| :
Db      52 LSPKTLLATNRFVCEICNKGFQRD---QNLQLH-RRGHNLPWKL--------------R 92
```

```
Qy    121 RKSSRTLKKG---------PRQDPS-AIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILK 170
          :||:: :||            |||  |: :  |:||        :||   |:       |
Db     93 QKSNKEVKKKVYVCPEVSCVHHDPSRALGDLTGIKKHFCRKHGEKKW--KCD----KCSK 146

Qy    171 LVAAQVGSQWKDIYQFLCNASEREV---AAFSNGYTADHERAYA-ALQHWTIRGPEASLA 226
            | |  | ||  :  :|   | :      ||   :    ||:  ||     |   :
Db    147 KYAVQ--SDWK-AHSKICGTKEYKCDCGTLFSRRDSFITHRAFCDALAEENARSHHS--- 200

Qy    227 QLISALRQH-----RRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLE 281
            : :|:     |:|  |                              |:|: |:|: : :||::
Db    201 ---QSKKQNPEILTRKNPV----------------------PNPV-PAPVDTESAKIK 232

Qy    282 NSALLTVEPSPQDK-------------------------------------------NKG 298
          :|: ||:: |   |                                          :|
Db    233 SSSTLTIKQSESPKTPPEIVQEAPKPTSLNVVTSNGVFAGLFESSSASPSIYTTSSSSKS 292

Qy    299 FFVDES--EPLLRCDSTSSGSSALSRN 323
            |   |  ||:    ||| ||| |  |
Db    293 LFASSSSIEPISLGLSTSHGSSFLGSN 319
```

RESULT 8
E71881
hypothetical protein jhp0842 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C;Accession: E71881
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557
A;Accession: E71881
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-524 <ARN>
A;Cross-references: GB:AE001513; GB:AE001439; NID:g4155407; PIDN:AAD06420.1; PID:g415541
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0842

Query Match 5.0%; Score 105.5; DB 2; Length 524;
Best Local Similarity 18.8%; Pred. No. 2.4;
Matches 93; Conservative 67; Mismatches 180; Indels 155; Gaps 19;

```
Qy      2 NSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHILK 61
          |:   || |    |  :  :    | :     |  :|| |       :: |:|   ||
Db     11 NANALNSGAKNEVKDAKNAPKSASKDFSKILNQKISKDKTAP-------KESPNH-SALK 62

Qy     62 LLP---SMEATGGEKSST-----------------------PIKGPKRGHPRQNLHKHF 94
            |     :|   ||: |                       |   |   |  |  |:
Db     63 ATPKDAKEDAKALEKTPTLPHQHAQNLAKDQQAPTLKDWLNHPKTHPTAPHEAQ--HETH 120

Qy     95 DINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKA------------ 142
          : ||  |                  :  : || |  :: |: :   |
Db    121 EANETNP------------------KTPNETLSKNEKK-PNEVASNAHQTNLPNKNPIT 160

Qy    143 ------GLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVA 196
                :|   ||| | :           |   |   |  ||  |:     |||  : |
Db    161 PNHANNAIKNPTTPTHNVK----------DPKTLKDIQTLSQKHDL-----NASNIQAA 204

Qy    197 AFSNGYTADHERAYAALQHWTIRGPEASL----------AQLISAL----------RQ 234
              | :::    |  |  ::  :||:           | | | |           ::
Db    205 T-----TPENKNPLNASDHLALKTTQASINHTLTKNDAKNTANLSSVLQSLEKKEPQNKE 259

Qy    235 HRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIP------SPNAKLENS----- 283
          |      ||    :::  |:   |    :|      :||       :|:|  ||:
Db    260 HANPPNNEKKTPPLKEALQMNAIKRDKTLSKKKPEKTPIHAKTQTIAPSATPENAPKIPL 319

Qy    284 ------ALLTVEPSPQD--------KNKGFFVDESEPLLRCDSTSSGSSALSRNGSFITK 329
                |:   | | |        : |   : :::   : :: ||  |  :   |  |
Db    320 KTLPLMPLIGANPPPNDNAPTPLEKEEKTKEISDNKEKAK-ETNSSAQSVQNTQASDKTS 378

Qy    330 EKKDTV--------LRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQAEDKLDRLFEI 381
          : | |          :|:: :  : :|    :   | |:||  :|   |   |: :  ::
Db    379 DNKSTTPKETIRHFTQQLKQEIQEYKPPMSKISMDLFPKELGKVEVTIQ---KVGKNLKV 435

Qy    382 IGVKSQEASQTLLDS 396
            :    : || ||:
Db    436 SVISHNNSLQTFLDN 450
```

RESULT 9
T04426
hypothetical protein T18B16.20 - Arabidopsis thaliana
N;Alternate names: hypothetical protein F13C5.220
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
C;Accession: T04426; T05042
R;Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansorge, W.; Bancroft, I.; Mewes, submitted to the Protein Sequence Database, April 1998
A;Reference number: Z15359
A;Accession: T04426
A;Molecule type: DNA
A;Residues: 1-1405 <BEV>
A;Cross-references: EMBL:AL021687
A;Experimental source: cultivar Columbia; BAC clone T18B16
R;Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Sch submitted to the Protein Sequence Database, February 1998
A;Reference number: Z15395
A;Accession: T05042
A;Molecule type: DNA
A;Residues: 1-1405 <BEW>
A;Cross-references: EMBL:AL021711
A;Experimental source: cultivar Columbia; BAC clone F13C5
C;Genetics:
A;Map position: 4
A;Introns: 893/3; 1164/2; 1183/3; 1192/3; 1198/2; 1214/3; 1251/1; 1282/2; 1327/1; 138
A;Note: T18B16.20; F13C5.220

Query Match 5.0%; Score 105.5; DB 2; Length 1405;
Best Local Similarity 22.1%; Pred. No. 9.5;
Matches 93; Conservative 69; Mismatches 156; Indels 103; Gaps 21;

```
Qy     17 LSSIQEGTVPDNTSSARGKEDV-------NKTLPNLQVVNHQQGPHHRHILKLLPSMEAT 69
          ||      :||  |     :::        ||||||: : :             |   : :
Db    756 LSETNLSELPDKISELSNLKELIIRKCSKLKTLPNLEKLTN-----------LEIFDVS 803

Qy     70 GGEKSSTPIKGPKRGHPRQNLHKHFDINE----HLPWMIVLFLLLVLVVIVVCSIRKSSR 125
          |  :  | |:|         |||  :::|     ||  |     |  :::  ||  |:
Db    804 GCTELET-IEGSFEN--LSCLHK-VNLSETNLGELPNKISELSNLKELILRNCSKLKALP 859

Qy    126 TLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVG----SQWK 181
           |:|        | : :|        :: |   | |        ::| |        ||
Db    860 NLEK---LTHLVIFDVSGCTNLDKIEESFESMSYLCESILCSSKRIVLADSSCIERDQWS 916

Qy    182 DIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVV 241
          | :  |  |: | ::|||      |:       : :  ||  |                |:|
Db    917 QIKE--CLTSKSEGSSFSNVGEKTREKLLYHGNRYRVIDPEVPL----------NIDIV 963

Qy    242 EKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFV 301
          :     :: :| |:|: :|     | :: :   | ::  :   : :|        || :|
Db    964 D-----IKRSTDLKTEYIAKAEYVS-IAENGSKSVSSLFDELQMASV--------KGCWV 1009

Qy    302 D---------ESEPLLRCDSTSSGS------------SALSRNGSFITKEKKDTVLRQV 339
          :          ||:  |  : :|| |              |  |  | || |      |:::
Db   1010 ERCKNMDVLFESDEQLEKEKSSSPSLQTLWISNLPLLTSLYSSKGGFIFKN-----LKKL 1064

Qy    340 RLDPC-DLQPIFDDMLHFLNPEELRVIEEIPQAEDKLDRLFEIIGVKSQEASQT----LL 394
          :| |  :: :| ::     |: | ::    :  |||:||||   ||: | |:     ||
Db   1065 SVDCCPSIKWLFPEI-----PDNLEILR--VKFCDKLERLFE---VKAGELSKLRKLHLL 1114
```

```
Qy    395 D 395
          |
Db   1115 D 1115
```

RESULT 10
T40906
probable ATP-dependent DNA helicase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: T40906
R;Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21956
A;Accession: T40906
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-607 <MUR>
A;Cross-references: EMBL:AL034490; PIDN:CAA22471.1; GSPDB:GN00068; SPDB:SPCC126.02c
A;Experimental source: strain 972h-; cosmid c126
C;Genetics:
A;Gene: SPDB:SPCC126.02c
A;Map position: 3
A;Introns: 11/3; 72/3; 105/1; 501/3; 571/3
C;Superfamily: thyroid autoantigen

Query Match 5.0%; Score 104.5; DB 2; Length 607;
Best Local Similarity 20.1%; Pred. No. 3.5;
Matches 91; Conservative 73; Mismatches 158; Indels 131; Gaps 24;

```
Qy      34 GKEDVNKTLPNLQVVNHQQGPHHRHILKLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKH 93
           |::||:       :||  |  :|:| ::            | :: ||| |     |
Db     204 GRQDVS------NLVNRGQA-QLQHMLNMI-----------TALQKPKRAH----FHLK 240

Qy      94 FDINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPR---QDPSAIVEKAGLKKSMTP 150
            |:   :   :  |:||  :       |::   || |     | :       || :
Db     241 MDLGNDVRIGVEAFILLKRL-----ESAKTNWVYAKGERFAVAVPQSKQVSFATKKELKK 295

Qy     151 TQNREKWIYYCNGHGI--------------DILKLVAAQVGSQWKDIY----QFLCNAS 191
           : |  : |   |  :                |::  :  |  |  :
Db     296 DEIRRSYSY--GGSSVVFGSDELNKVRSFEPPTLRIIGFRDFSTLKPWHCLKPAVFLRPK 353

Qy     192 EREVAAFSNGYTADHERAYAALQ---HWTIRGPEAS---LAQLISALRQHRRNDVVEKIR 245
           : |:      ::| |::  |: :    | :  | |:   :| | :    | |:|  |
Db     354 DDEIIGSGAVFSAIHKKLLASNKIGIAWFVSRPNANPCFVAMLATPGSIHIRDD-FELPL 412

Qy     246 GLMEDTTQLET--DKLALPMSPSPLSPSPIPSPNAKLEN---------------------- 282
           |:     || |  |  :||    |::|:||  |:  :|
Db     413 GIF--LVQLPTADDIRSLP----PINPNPISMPSNLIETMQRILRGMELRSYQPGKYNNP 466

Qy     283 ---------SALLTVEPSPQDKNKGFFVDESEPLLRCDSTSSG-------------SSA 319
                    ||   |  | |     ||| : |  :       |               :
Db     467 SLQWHYKVLQALALDEEIPTD-----FVDNTLPKYKAIQKRVGEYMGDVNNIVAEYRNDI 521

Qy     320 LSRNGSFITKEKKDTVLRQVRLDPCDLQPIF--DDMLHFLNPE-----ELRVIEEIPQAE 372
             :||    :| :  :::: |::    :|||  || |  |  |     |:: : :: | :
Db     522 SDKNGIKEEEEDQGPIVKKARIEKSG-KPIFAEDDRLKQLYIEGVLDKEIKAL-KVSQLK 579

Qy     373 DKL-DRLFEIIGVKSQEASQTLLDSVYSHLPDL 404
           | | ||   : | |:      |||:: :::  |
Db     580 DILRDRGLRVSGKKAD-----LLDNLTNYVKKL 607

RESULT 11
S47508
genome polyprotein - plum pox virus (isolate PVV-SC)
N;Contains: coat protein; protein 6K1; protein 6K2; protein CI; protein HCpro; protein N
C;Species: plum pox virus, PPV
A;Variety: isolate PVV-SC
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
C;Accession: S47508
R;Maiss, E.; Deborre, G.; Jelkmann, W.; Casper, R.
submitted to the EMBL Data Library, August 1994
A;Description: Complete nucleotide sequence of a plum pox potyvirus isolate (PVV-SC)
A;Reference number: S47508
A;Accession: S47508
A;Molecule type: genomic RNA
A;Residues: 1-3140 <MAI>
A;Cross-references: EMBL:X81083; NID:g531731; PIDN:CAA56974.1; PID:g531732
A;Experimental source: isolate PVV-SC
C;Superfamily: tobacco etch virus genome polyprotein
C;Keywords: ATP; coat protein; genome-linked protein; nucleotide binding; P-loop; pho
F;1-308/Product: protein P1 #status predicted <P1P>
F;309-766/Product: protein HCpro #status predicted <HCP>
F;767-1116/Product: protein P3 #status predicted <P3P>
F;1117-1168/Product: protein 6K1 #status predicted <6K1>
F;1169-1803/Product: protein CI #status predicted <CIP>
F;1253-1260/Region: nucleotide-binding motif A (P-loop)
F;1338-1343/Region: nucleotide-binding motif B
F;1342-1345/Region: DEXH motif
F;1804-1856/Product: protein 6K2 #status predicted <6K2>
F;1857-2292/Product: protein NIa #status predicted <NIA>
F;1857-2049/Product: VPg protein #status predicted <VPG>
F;2293-2810/Product: protein NIb #status predicted <NIB>
F;2811-3140/Product: coat protein #status predicted <COA>
F;1919/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Query Match 5.0%; Score 104.5; DB 2; Length 3140;
Best Local Similarity 20.7%; Pred. No. 34;
Matches 65; Conservative 52; Mismatches 102; Indels 95; Gaps 15;

```
Qy      92 KHFDINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKK--GPRQDP----------SAIV 139
           | ||       |:::| ||: :     |:    : ||:    |:|           : :
Db    1024 KMFD-------MLIVFGLLLSIGATCNSMVNEHKHLKQLAADREDKKRFKRLQVLHTRLS 1076

Qy     140 EKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFS 199
           || |     ||| |  ::: |  |   |:||     :|       | : : |:|:
Db    1077 EKVG----CTPTAN--EFLEYVGGENPDLLKHAEDLIGDG-----QVVVHQSKRD----- 1120

Qy     200 NGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND----VVEKIRGLMEDTTQ-- 253
             |: ||  |              |:  |    |:|    :: |::|:|    |
Db    1121 --SQANLERVVA-----------FVALVMMLFDSERSDGVYKILNKLKGIMGSVEQAV 1165

Qy     254 -----------LETDKLALP--MSPSPLSPS-PIPSPNAKLENSALLTVEPSPQDKNKGF 299
                      |:  || :   :  : ::|: |  |   |   : | |    |  : :|:
Db    1166 HHQSLDDIEDILDEKKLTVDFVLQSNEVAPTVPFDSTFEKWWTNQLETGNVIPHYRTEGY 1225

Qy     300 FVD---------------ESEPLLRCDSTSSGSSA-----LSRNGSFI----TKEKKDT 334
           |::                 : :|   :  || |      ||: |  :    |:   :
Db    1226 FLEFTRENAAHIANEVMHGSHQDILIRGAVGSGKSTGLPFHLSKKGHVLLIEPTRPLAEN 1285

Qy     335 VLRQVRLDPCDLQP 348
           | :|:|  | :: |
Db    1286 VCKQLRGQPFNVNP 1299
```

RESULT 12
S51323
SAC3 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YD8358.13; protein YDR159w
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 21-Jul-2000
C;Accession: S51323; S57983; S71744
R;Bauer, A.; Koelling, R.
submitted to the EMBL Data Library, January 1995
A;Description: The SAC3 gene codes for a nuclear protein required for normal mitosis.
A;Reference number: S51323
A;Accession: S51323
A;Molecule type: DNA
A;Residues: 1-1301 <BAU>
A;Cross-references: EMBL:Z47805; NID:g634085; PID:g634086
R;Murphy, L.; Richards, C.; Harris, D.
submitted to the EMBL Data Library, July 1995
A;Reference number: S57971
```

```
A;Accession: S57983
A;Molecule type: DNA
A;Residues: 1-1301 <MUR>
A;Cross-references: EMBL:Z50046; NID:g899393; PID:g899406; MIPS:YDR159w
A;Experimental source: strain AB972
R;Bauer, A.; Koelling, R.
Yeast 12, 965-975, 1996
A;Title: Characterization of the SAC3 gene of Saccharomyces cerevisiae.
A;Reference number: S71744; MUID:97027306
A;Accession: S71744
A;Molecule type: DNA
A;Residues: 1-1301 <BAW>
A;Cross-references: EMBL:Z47805; NID:g634085; PIDN:CAA87767.1; PID:g634086
C;Genetics:
A;Gene: SGD:SAC3; LEP1
A;Cross-references: SGD:S0002566; MIPS:YDR159w
A;Map position: 4R
C;Function:
A;Description: potential regulator of leucine permease gene expression
C;Keywords: nucleus; transmembrane protein
F;999-1015/Domain: transmembrane #status predicted <TMM>

  Query Match             4.9%;  Score 103.5;  DB 2;  Length 1301;
  Best Local Similarity   21.8%;  Pred. No. 12;
  Matches   69;  Conservative   46;  Mismatches  124;  Indels   77;  Gaps   12;

Qy     100 LPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIY 159
           || ::  || : |  |   ::|  | || |  :  |   :|     :|    ||:: |
Db     438 LPLLMGFFLQMHLTDIRFYALRALSHTLNKKHKPIPFIYLE------NMLLFNNRQEIIE 491

Qy     160 YCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWT-- 217
           :|| : |:|:   || :    | :  :    || :       |    ||    ||  |
Db     492 FCNYYSIEIINGDAADL----KTLQHYSHKLSETQ--PLKKTYLTCLER---RLQKTTYK 542

Qy     218 --IRGPEASLAQLI---------------SALRQHRRNDVVEKIRGLMEDTTQLETDKL 259
             | | | :||  :                | | :: :|:  ||:        |: |
Db     543 GLINGGEDNLASSVYVKDPKKDRIPSIADQSFLMENFQNNYNEKLNQNSSVKPQINTSPK 602

Qy     260 ALPMSPS--PLSPSPIPSPNAKLENSALLTVEP--SPQ----------------DKNKG 298
           :   |:  | |      |  ::  |  |:  |  :||                |
Db     603 RVATRPNHFPFSQESKQLP--QISQSHTLSTNPLLTPQVHGDLSEQKQQQIKTVTDGGSP 660

Qy     299 FFVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLN 358
           |  |:|     ::: :   : : ||::   ||  :   ::|
Db     661 FVFDQSAQNSTVEASKAHMISTTSNGAY--DEKLSSEQEEMR----------------K 701

Qy     359 PEELRVIEEIPQAEDK 374
           || |: ||  | : |
Db     702 KEEQRIEEEKTQLKKK 717


RESULT  13
T18532
serine/threoine protein kinase - guinea pig
N;Alternate names: STE20-like kinase
C;Species: Cavia porcellus (guinea pig)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C;Accession: T18532
R;Itoh, S.; Kameda, Y.; Yamada, E.; Tsujikawa, K.; Mimura, T.; Kohama, Y.
Arch. Biochem. Biophys. 340, 201-207, 1997
A;Title: Molecular cloning and characterization of a novel putative STE20-like kinase in
A;Reference number: Z18952; MUID:97288344
A;Accession: T18532
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1231 <ITO>
A;Cross-references: EMBL:D88425; NID:g2911113; PIDN:BAA24930.1; PID:g2911114
A;Experimental source: strain Hartley; liver

  Query Match             4.9%;  Score 103;  DB 2;  Length 1231;
  Best Local Similarity   25.1%;  Pred. No. 12;
  Matches   53;  Conservative   27;  Mismatches   75;  Indels   56;  Gaps   10;

Qy     231 ALRQHRRNDVVEKIRGLMEDTTQLETDK-------------LALPMSPSPLSPSPIPSP 276
           |: :    ||  : ||     ||::|:|              : :|  ||   |:|||
Db     650 AVSETDEEDVQSETRGAPMAVTQMDTEKNETPHEAPAQVEVQVPVPPQPSEPPPAPIPSI 709

Qy     277 N------------AKLENSALLTVEPSPQ-----DKNKGFFVDESEPLLRCDSTSSGSS 318
           |             | |    :|  |   |     |   |  | |      |  | ||
Db     710 NINSEAAENKGEMGASLNTETILLPESESQKENDTDSGTGSTADNSS----IDLNLSISS 765

Qy     319 ALSR---NGSFI---TKEKKDTV--LRQVRLDPCDL----QPIFDDM------LHFLNPE 360
            ||:   |||     |: :| |:   |:  :|  ::      |  |       | ||  :
Db     766 FLSKTKDNGSISLQETRRQKKTLKKTRKFIVDGVEVSVTTSKIVTDSDSKTEELRFLRRQ 825

Qy     361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQ 391
           ||| :  : : | :  :  :: |   |:  |
Db     826 ELRELRFLQKEEQRAQQ--QLNGKLQQQREQ 854


RESULT  14
T13806
toucan gene protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13806
R;Grammont, M.; Dastugue, B.; Couderc, J.L.
Development 124, 4917-4926, 1997
A;Title: The Drosophila toucan (toc) gene is required in germline cells for somatic c
A;Reference number: Z17769; MUID:98090047
A;Accession: T13806
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2176 <GRA>
A;Cross-references: EMBL:Y14157; NID:g2760521; PIDN:CAA74574.1; PID:g2760522
C;Genetics:
A;Gene: toc
A;Cross-references: FlyBase:FBgn0015600

  Query Match             4.9%;  Score 103;  DB 2;  Length 2176;
  Best Local Similarity   22.1%;  Pred. No. 27;
  Matches   96;  Conservative   60;  Mismatches  170;  Indels  108;  Gaps   25;

Qy       3 STESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTL--PNLQVVNHQQGPHHRHIL 60
           | |:: :  |    |:|:     || |:    |  :| ||    |  ::  | | |  :|
Db     467 SQETDDTLDV---TLTSL----APDKTNM---KLPLNSTLNTERLLDISALQSPRHIQLL 516

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNL-----------HKHFDINEHLPWMIVLFL 108
           |   :|   |  |   ::  :|  |: :|            |    : :  |  : |
Db     517 NLTQELER--GTPSRVHLQSGRRSMPQHSLLCLSPQSATTTPHGMRMMGQATPQPVHLLS 574

Qy     109 LLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAG---LKKSMTPTQNREKWIYYCNGHG 165
           |:          | :::    ] : |   :   |   |  ::| |  | :  |    |
Db     575 PLL----------KQAQSALVLPTRTPDGDITMDGNGALDSTLTSTSVRGRTRY---SFG 621

Qy     166 IDI--------LKLVAAQVGS--QWKDI-YQFLCNASEREVAAFSNGYTADHERAYAALQ 214
           :|:         ::||      |  | : :  | |   |  ::       | |  : :
Db     622 LDLPDTTLDCSIELVDNSFSSSTQLQQLQQQLLKKQSSFDLDESLGILTPDQMKDF---- 677

Qy     215 HWTIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPI- 273
             :  |  :|   :  :| |  |  : ::|  || |    |     ||: |  :    :
Db     678 ---LDSPHNNLMHNLEMIRMHHPN--LLQLR--MEQTPSPE----ELPLDPIEIKSEIVK 726

Qy     274 ----------PSPNAKLENSALLTVEPSPQDKNKGFFVD--ESEPLLR--CDSTSSGSSA 319
                     | ::|| || : :|  |     : |:  |   | |  |  || : |
Db     727 QVEASQNQVNTSQHSKLSNSFITSV-TSVTSLDTGYQGDGEMSRPASRGACDHSPSNGPH 785

Qy     320 L---SRNGSFITKEKKDTVLRQVRLDPCDLQPIF-----DDMLH-----------FLNP 359
           |   ||  ||         ||  | ||       |     ||:||              |
Db     786 LGRVSRQPSF--PPPNPAPLR--RQDPMTDSDFFTESDADDVLHRGDRRAQVIDGQLYGP 841

Qy     360 EELRVIEEIPQAED 373
           : ::    :|| ||
```

```
Db    842 DMMQPSASVPQMED 855


RESULT  15
T28067
hypothetical protein ZK867.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28067
R;Nhan, M.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid ZK867.
A;Reference number: Z20464
A;Accession: T28067
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-494 <NHA>
A;Cross-references: EMBL:U41039; PIDN:AAA82395.1; CESP:ZK867.1
C;Genetics:
A;Gene: CESP:ZK867.1
A;Introns: 3/3; 46/1; 203/2; 242/1; 265/3; 339/1; 402/3


  Query Match             4.9%;  Score 101.5;  DB 2;  Length 494;
  Best Local Similarity   19.2%;  Pred. No. 4.5;
  Matches   87;  Conservative   57;  Mismatches  141;  Indels  167;  Gaps   22;

Qy     30 SSARGKED-VNKTLPNLQVVNHQQGPHHRHILKLLPSMEATGGEKSSTPIKGP------- 81
          : |  |:| : |  |  ::|    ||              | |   :|    |
Db     61 NEAIAKDDIIVKKAP--KIVTKDNGP------------TTNGSTPTTSTATPSVITVSS 105

Qy     82 ----KRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLK---KGP--- 131
               ||   |       | |                  ::  : ||:|   : |
Db    106 ALASSNGHNNNN-------NNH------------------AVNNNLRTIKMELEDPDYN 139

Qy    132 ---RQDPSAIVEKAGLKKSMTP-TQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFL 187
             :  |: :| |     ::|| ::   | |        :||: :|  ||
Db    140 LIAKSAPTPVVSKIVATHTVTPRSRPTPKDIK-------EILETIAPSVG---------- 182

Qy    188 CNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLIS---------ALRQHRRN 238
          : :  |:       ::: :|:                  |||         :|   :
Db    183 VSETPEEMCLLPKDASSESDRSV----------------LISLGFDFGSTLSLNHQQLQ 225

Qy    239 DVVEKIRG---LMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTV------- 288
          || :::|    :  || | :         || |:: : | :   :   :|::
Db    226 QVVRELKGELSISPDTVQSDHSDDFEQDSPPPMAIANISTVGGEATLAAMIVAATNASGQ 285

Qy    289 ----EPSPQDKNKGF-----FVDESEPLLRCDSTSSGSSALS--------RNGSFITKEK 331
              |   |  ||         ||:|     |||| |  |           |: : |
Db    286 RGDGTPDSTDTQKGCSPQRELSPESDP-----STSSGDSCPSPPKMLHCKECGTLVRKSS 340

Qy    332 ------------KDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEE----------- 367
                      |   |   |   ||:    ||     |||||
Db    341 HLPIHMTMSHGYPPPLVAAPVEEKPAPEQPVNASSLH----NELRVISNAICELKAQQAA 396

Qy    368 IPQAEDKLDRLFEIIGVKSQEASQTLLDSVYS 399
          |: |  |  :   :| | : : :| |:|:|:
Db    397 TPRVEQALTYIDSRVG-KLERSLETALNSIYT 427


Search completed: January  4, 2002, 08:38:38
Job time: 121 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:37:27 ; Search time 13.26 Seconds
(without alignments)
1119.853 Million cell updates/sec

Title: US-09-733-956-2
Perfect score: 2091
Sequence: 1 MNSTESNSSASVRPKVLSSI.....SQEASQTLLDSVYSHLPDLL 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	103.5	4.9	1301	1	SAC3_YEAST	P46674	saccharomyc
2	101	4.8	342	1	TONB_PSEAE	Q51368	pseudomonas
3	100	4.8	746	1	DTNA_MOUSE	Q9d2n4	mus musculu
4	98.5	4.7	3140	1	POLG_PPVRA	P17767	p genome po
5	98	4.7	897	1	APG1_YEAST	P53104	saccharomyc
6	97	4.6	1971	1	MC3A_MOUSE	Q9wuu9	mus musculu
7	96	4.6	969	1	YEAC_SCHPO	O14077	schizosacch
8	95.5	4.6	733	1	RRP6_YEAST	Q12149	saccharomyc
9	95.5	4.6	1025	1	MK21_YEAST	Q12176	saccharomyc
10	95.5	4.6	1483	1	CYP1_YEAST	P12351	saccharomyc
11	95.5	4.6	3140	1	POLG_PPVSK	Q84934	p genome po
12	95	4.5	707	1	YJ9C_YEAST	P47166	saccharomyc
13	95	4.5	3141	1	POLG_PPVD	P13529	p genome po
14	94	4.5	249	1	YCX9_CHLVU	O20173	chlorella v
15	94	4.5	994	1	CLC1_MOUSE	Q64347	mus musculu
16	93.5	4.5	1108	1	DBS_HUMAN	O15068	homo sapien
17	93	4.4	465	1	SPE4_CAEEL	Q01608	caenorhabdi
18	93	4.4	470	1	SELA_MOOTH	O33277	moorella th
19	93	4.4	723	1	GGA3_HUMAN	Q9nz52	homo sapien
20	92.5	4.4	910	1	YK69_YEAST	P36165	saccharomyc
21	92.5	4.4	2468	1	MAPB_HUMAN	P46821	homo sapien
22	92	4.4	994	1	SYLM_NEUCR	P15181	neurospora
23	91.5	4.4	627	1	Y017_RICPR	Q9zec6	rickettsia
24	91	4.4	511	1	CP4B_RAT	P15129	rattus norv
25	91	4.4	1770	1	RI15_YEAST	P43565	saccharomyc
26	91	4.4	2109	1	PKS1_ASPPA	Q12053	aspergillus
27	90.5	4.3	760	1	FXM1_MOUSE	O08696	mus musculu
28	90	4.3	712	1	RR3_CHLRE	Q08365	chlamydomon
29	90	4.3	1443	1	E75C_DROME	P13055	drosophila
30	89.5	4.3	593	1	BIN1_HUMAN	O00499	homo sapien
31	89.5	4.3	1068	1	AF10_MOUSE	O54826	mus musculu
32	89.5	4.3	1637	1	MRSP_STAAU	P80544	staphylococ
33	89	4.3	409	1	YNP2_YEAST	P53901	saccharomyc
34	89	4.3	5430	1	ACF7_HUMAN	Q9upn3	homo sapien
35	88.5	4.2	1332	1	SPT7_YEAST	P35177	saccharomyc
36	88.5	4.2	1390	1	RPOB_MYCGA	P47715	mycoplasma
37	88	4.2	501	1	C72N_ARATH	Q9ltm0	arabidopsis
38	88	4.2	688	1	PYS2_PSEAE	Q06584	pseudomonas
39	88	4.2	836	1	Y167_HUMAN	Q99490	homo sapien
40	88	4.2	939	1	ST20_YEAST	Q03497	saccharomyc
41	88	4.2	1176	1	NIR_NEUCR	P38681	neurospora
42	87.5	4.2	622	1	YNA3_CAEEL	P45894	caenorhabdi
43	87.5	4.2	688	1	EFG_MYCPN	P75544	mycoplasma
44	87.5	4.2	736	1	GEPH_RAT	Q03555	rattus norv
45	87.5	4.2	914	1	PBPA_BACSU	P39793	bacillus su

ALIGNMENTS

RESULT 1
SAC3_YEAST
ID SAC3_YEAST STANDARD; PRT; 1301 AA.
AC P46674;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LEUCINE PERMEASE TRANSCRIPTIONAL REGULATOR.
GN SAC3 OR LEP1 OR YDR159W OR YD8358.13.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97027306; PubMed=8873450;
RA Bauer A., Koelling R.;
RT "Characterization of the SAC3 gene of Saccharomyces cerevisiae.";
RL Yeast 12:965-975(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Murphy L., Richards C., Harris D., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 189-1301 FROM N.A.
RC STRAIN=S288C / DBY939;
RA Stella C.A., Korch C., Ramos E.H., Mattoon J.R.;
RT "Cloning and sequencing of LEP1, a gene associated with leucine
RT transport.";
RL Yeast 11:460-460(1995).
CC -!- FUNCTION: POTENTIAL REGULATOR OF LEUCINE PERMEASE GENE(S)
CC EXPRESSION. REQUIRED FOR NORMAL MITOSIS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SAC3 FAMILY.
CC ---

CC ---
DR EMBL; Z47805; CAA87767.1; -.
DR EMBL; Z50046; CAA90379.1; -.
DR EMBL; U35227; AAA79056.1; -.
DR SGD; S0002566; SAC3.
KW Transcription regulation; Nuclear protein.
SQ SEQUENCE 1301 AA; 149568 MW; 0679DB1673DDACEB CRC64;

Query Match 4.9%; Score 103.5; DB 1; Length 1301;
Best Local Similarity 21.8%; Pred. No. 6.1;
Matches 69; Conservative 46; Mismatches 124; Indels 77; Gaps 12;

```
Qy     100 LPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIY 159
           || ::  || : |  |   ::|  | || |  :  |   :|      :|    ||:: |
Db     438 LPLLMGFFLQMHLTDIRFYALRALSHTLNKKHKPIPFIYLE------NMLLFNNRQEIIE 491

Qy     160 YCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWT-- 217
           :|| : |:|:   || :   | : :   || :      |    ||    ||  |
Db     492 FCNYYSIEIINGDAADL----KTLQHYSHKLSETQ--PLKKTYLTCLER---RLQKTTYK 542

Qy     218 --IRGPEASLAQLI---------------SALRQHRRNDVVEKIRGLMEDTTQLETDKL 259
             | | | :||  :                | | :: :|:  ||:        |: |
Db     543 GLINGGEDNLASSVYVKDPKKDRIPSIADQSFLMENFQNNYNEKLNQNSSVKPQINTSPK 602

Qy     260 ALPMSPS--PLSPSPIPSPNAKLENSALLTVEP--SPQ----------------DKNKG 298
           :   |:  | |      |  ::  |  |:  |  :||                |
Db     603 RVATRPNHFPFSQESKQLP--QISQSHTLSTNPLLTPQVHGDLSEQKQQQIKTVTDGGSP 660

Qy     299 FFVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLN 358
           |  |:|      ::: :   : :||::   ||  :   ::|
Db     661 FVFDQSAQNSTVEASKAHMISTTSNGAY--DEKLSSEQEEMR----------------K 701

Qy     359 PEELRVIEEIPQAEDK 374
           || |: ||  | : |
Db     702 KEEQRIEEEKTQLKKK 717
```

```
RESULT   2
TONB_PSEAE
ID   TONB_PSEAE     STANDARD;      PRT;   342 AA.
AC   Q51368;
DT   01-NOV-1997 (Rel. 35, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   TONB PROTEIN.
GN   TONB OR PA5531.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=96262716; PubMed=8704984;
RA   Poole K., Zhao Q., Neshat S., Heinrichs D.E., Dean C.R.;
RT   "The Pseudomonas aeruginosa tonB gene encodes a novel TonB protein.";
RL   Microbiology 142:1449-1458(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- FUNCTION: INTERACTS WITH OUTER MEMBRANE RECEPTOR PROTEINS THAT
CC       CARRY OUT HIGH-AFFINITY BINDING AND ENERGY DEPENDENT UPTAKE INTO
CC       THE PERIPLASMIC SPACE OF SPECIFIC SUBSTRATES. IT COULD ACT TO
CC       TRANSDUCE ENERGY FROM THE CYTOPLASMIC MEMBRANE TO SPECIFIC ENERGY-
CC       REQUIRING PROCESSES IN THE OUTER MEMBRANE, RESULTING IN THE
CC       RELEASE INTO THE PERIPLASM OF LIGANDS BOUND BY THESE OUTER
CC       MEMBRANE PROTEINS (BY SIMILARITY).
CC   -!- SUBUNIT: THE ACCESSORY PROTEINS EXBB AND EXBD SEEM TO FORM A
CC       COMPLEX WITH TONB (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC. ANCHORED TO THE CYTOPLASMIC
CC       MEMBRANE VIA ITS N-TERMINAL SIGNAL-LIKE SEQUENCE, SPANS THE
CC       PERIPLASM (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TONB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U23764; AAB18654.1; -.
DR   EMBL; AE004965; AAG08916.1; -.
DR   InterPro; IPR002965; P_rich_extensn.
DR   InterPro; IPR003538; TonB.
KW   Transport; Protein transport; Inner membrane; Periplasmic;
KW   Transmembrane; Signal-anchor; Repeat; Complete proteome.
FT   DOMAIN        1     87       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     88    108       SIGNAL-ANCHOR (POTENTIAL).
FT   DOMAIN      109    342       PERIPLASMIC (POTENTIAL).
FT   DOMAIN      159    181       9 X 2 REPEATS OF K-P.
FT   DOMAIN      190    199       4 X 2 REPEATS OF K-P.
FT   CONFLICT    272    272       G -> R (IN REF. 1).
SQ   SEQUENCE   342 AA;  36892 MW;  157F036B39E9152D CRC64;
```

```
  Query Match             4.8%;  Score 101;  DB 1;  Length 342;
  Best Local Similarity   20.0%;  Pred. No. 1.5;
  Matches   73;  Conservative   40;  Mismatches  122;  Indels  130;  Gaps   17;

Qy      64 PSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRKS 123
           |:  | | |  | |    :|| |  ||                          :| ||
Db      22 PTAPAQGDESESLPCVNAQRGEP--NLR----------------------VVDCS---- 52

Qy     124 SRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDI 183
                 | |:|    ||              |  | |   || |   :      |:|
Db      53 ------GARRDEEVAVE-------------EVLIPY--AHGSDPEDVPGEPPKSRW--- 87

Qy     184 YQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGP--------EASLAQLISALRQH 235
                       |    | |     ||  | :  |        |      :: ::
Db      88 ------------WLSSGAAVAMHVAIIGALV-WVMPTPAELNLGHGELPKTMQVNFVQLE 134

Qy     236 RRNDVVEKIRGLMEDTT--QLETDKLALPMSPSPL-SPSPIPSPNAKLENSALLTVEPSP 292
           :: :   :       : |  ::|  |   |  | |:  | | | |  |   :|:   :| |
Db     135 KKAEPTPQPPAAAPEPTPPKIEEPK-PEPPKPKPVEKPKPKPKPKPKPVENAIPKAKPKP 193

Qy     293 QDKNKGFFVDESEPLLRCDSTSSGSSA----------------LSRNGS-----FITKEK 331
           : | |     | ||     |  | |||                 : :||       :
Db     194 EPKPK----PEPEPSTEASSQPSPSSAAPPPAPTVGQSTPGAQTAPSGSQGPAGLPSGSL 249

Qy     332 KDTVLRQVRLDPCDLQPIFDDMLHFLNPE----------------ELRVIEEIPQAEDKL 375
           |: :: :|:||    |::  |       |                 :::|:| :|
Db     250 NDSDIKPLRMDP----PVYPRMAQARGIEGRVKVLFTITSDGRIDDIQVLESVP------ 299

Qy     376 DRLFE 380
           |:|:
Db     300 SRMFD 304
```

```
RESULT   3
DTNA_MOUSE
ID   DTNA_MOUSE     STANDARD;      PRT;   746 AA.
AC   Q9D2N4; Q61498; Q61499; Q9WUM0; Q9WUL9; P97319; Q9QZZ5;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   DYSTROBREVIN ALPHA (ALPHA-DYSTROBREVIN).
GN   DTNA OR DTN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 2 AND 7).
RC   TISSUE=Brain;
RX   MEDLINE=96205975; PubMed=8631824;
```

```
RA   Blake D.J., Nawrotzki R., Peters M.F., Froehner S.C., Davies K.E.;
RT   "Isoform diversity of dystrobrevin, the murine 87-kDa postsynaptic
RT   protein.";
RL   J. Biol. Chem. 271:7802-7810(1996).
RN   [2]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   STRAIN=129/SV; TISSUE=Brain;
RX   MEDLINE=97224465; PubMed=9119373;
RA   Ambrose H.J., Blake D.J., Nawrotzki R.A., Davies K.E.;
RT   "Genomic organization of the mouse dystrobrevin gene: comparative
RT   analysis with the dystrophin gene.";
RL   Genomics 39:359-369(1997).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 5).
RC   TISSUE=Muscle;
RX   MEDLINE=98369056; PubMed=9701558;
RA   Nawrotzki R., Loh N.Y., Ruegg M.A., Davies K.E., Blake D.J.;
RT   "Characterisation of alpha-dystrobrevin in muscle.";
RL   J. Cell Sci. 111:2595-2605(1998).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 3; 4 AND 6).
RC   STRAIN=C3H; TISSUE=Muscle;
RX   MEDLINE=20035753; PubMed=10570976;
RA   Enigk R.E., Maimone M.M.;
RT   "Differential expression and developmental regulation of a novel
RT   alpha-dystrobrevin isoform in muscle.";
RL   Gene 238:479-488(1999).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=C57BL/6J; TISSUE=Skin;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: INVOLVED IN SYNAPSE MATURATION AND REQUIRED FOR NORMAL
CC       MUSCLE FUNCTION.
CC   -!- SUBUNIT: INTERACTS WITH DYSTROPHIN, UTROPHIN AND SYNTROPHIN.
CC       ISOFORMS 5 AND 6 DO NOT INTERACT WITH SYNTROPHIN. ISOFORMS 3 AND 4
CC       DO NOT INTERACT WITH UTROPHIN.
CC   -!- SUBCELLULAR LOCATION: CYSTOPLASMIC.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 7 ISOFORMS; 1 (SHOWN HERE),
CC       2/ALPHA-DYSTROBREVIN-1/ALPHA-DB1, 3/ALPHA-DYSTROBREVIN-2B/ALPHA-
CC       DB2B, 4/ALPHA-DYSTROBREVIN-2A/ALPHA-DB2A, 5/ALPHA-DYSTROBREVIN-
CC       3/ALPHA-DB3, 6/ALPHA-DYSTROBREVIN-3/ALPHA-DB3 AND 7; ARE PRODUCED
CC       BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL MUSCLE, HEART, LUNG AND
CC       BRAIN. SARCOLEMMA AND NEUROMUSCULAR JUNCTION IN SKELETAL MUSCLE.
CC       ISOFORM 2 IS RESTRICTED TO THE NEUROMUSCULAR JUNCTION. ISOFORMS 5
CC       AND 6 ARE ONLY EXPRESSED IN MUSCLE.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSION OF ALPHA-DYSTROBREVIN IS
CC       UPREGULATED DURING DIFFERENTIATION, WITH ISOFORMS 2, 5 AND 6
CC       EXPRESSED EARLIEST AND ISOFORM 3 AND 4 EXPRESSED LATER.
CC   -!- DOMAIN: THE COILED-COIL DOMAIN MEDIATE THE INTERACTION WITH
CC       DYSTROPHIN AND UTROPHIN.
CC   -!- PTM: PHOSPHORYLATION OF ISOFORM 2 ON TYROSINE KINASE SUBSTRATE
CC       DOMAIN PRESENT IN THE C-TERMINUS.
CC   -!- SIMILARITY: BELONGS TO THE DYSTROPHIN FAMILY. DYSTROBREVIN
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X95226; CAA64518.1; -.
DR   EMBL; X95227; CAA64519.1; -.
DR   EMBL; Z79787; CAB02145.1; -.
DR   EMBL; Z79788; CAB02145.1; JOINED.
DR   EMBL; Z79789; CAB02145.1; JOINED.
DR   EMBL; Z79790; CAB02145.1; JOINED.
DR   EMBL; Z79791; CAB02145.1; JOINED.
DR   EMBL; Z79792; CAB02145.1; JOINED.
DR   EMBL; Z79793; CAB02145.1; JOINED.
DR   EMBL; Z79794; CAB02145.1; JOINED.
DR   EMBL; Z79795; CAB02145.1; JOINED.
DR   EMBL; Z79796; CAB02145.1; JOINED.
DR   EMBL; Z79797; CAB02145.1; JOINED.
DR   EMBL; AJ009669; CAA08770.1; -.
DR   EMBL; AF143544; AAD33915.1; -.
DR   EMBL; AF143543; AAD33914.1; -.
DR   EMBL; AF143542; AAD33913.1; -.
DR   EMBL; AK019477; BAB31746.1; -.
DR   MGD; MGI:106039; Dtn.
DR   InterPro; IPR000345; CytC_heme_bind.
DR   InterPro; IPR000433; ZnF_ZZ.
DR   Pfam; PF00569; ZZ; 1.
DR   SMART; SM00291; ZnF_ZZ; 1.
KW   Synapse; Coiled coil; Zinc-finger; Phosphorylation;
KW   Alternative splicing.
FT   ZN_FING      237    284       ZZ-TYPE.
FT   DOMAIN       397    447       SYNTROPHIN-BINDING REGION.
FT   DOMAIN       458    557       COILED COIL (POTENTIAL).
FT   VARSPLIC     334    334       V -> VDTW (IN ISOFORM 5 AND ISOFORM 7).
FT   VARSPLIC     363    419       MISSING (IN ISOFORM 2 AND ISOFORM 7).
FT   VARSPLIC     364    371       PPKDSEVE -> DGAHGGCV (IN ISOFORM 5 AND
FT                                 ISOFORM 6).
FT   VARSPLIC     372    746       MISSING (IN ISOFORM 5 AND ISOFORM 6).
FT   VARSPLIC     559    567       TQGASSPRS -> VSYVPYCRS (IN ISOFORM 4 AND
FT                                 ISOFORM 7).
FT   VARSPLIC     568    746       MISSING (IN ISOFORM 4 AND ISOFORM 7).
FT   VARSPLIC     560    561       QG -> R (IN ISOFORM 2).
FT   VARSPLIC     559    570       TQGASSPRSSPS -> VMHEIIPLEERT (IN ISOFORM
FT                                 3).
FT   VARSPLIC     571    746       MISSING (IN ISOFORM 3).
FT   CONFLICT     241    241       V -> L (IN REF. 1; CAA64519).
FT   CONFLICT     251    251       I -> M (IN REF. 1, 2, 3 AND 4).
FT   CONFLICT     259    259       Q -> P (IN REF. 4).
FT   CONFLICT     416    416       D -> N (IN REF. 4).
SQ   SEQUENCE   746 AA;  84049 MW;  0F06282718E21679 CRC64;


  Query Match             4.8%;  Score 100;  DB 1;  Length 746;
  Best Local Similarity   25.1%;  Pred. No. 5.3;
  Matches   68;  Conservative   31;  Mismatches  102;  Indels   70;  Gaps   13;

Qy     132 RQDPSAIVEK----------------AGLKKSMTPTQNREKWIYYCNGHGIDILKLVA 173
           | ||  ::|                |    |  ||| |           ||   :
Db     414 RNDPPCMLESSNRLDEEHRLIARYAARLAAESSSSQPTQQRS---------APDISFTID 464

Qy     174 AQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERA-----YAALQHWT-------IRGP 221
           |    | : :   | |   ||:         :||:|       | |: |       :|
Db     465 A--NKQQRQLIAELEN-KNREILQEIQRLRVEHEQASQPTPEKAQQNPTLLAELRLLRQR 521

Qy     222 EASLAQLISALRQHRRNDVVEKIRGLM----EDTTQLETDKLALP-MSPSPLSPSPIPSP 276
           :  | | :|||:: || ::: :: |||    |:  :  |   : |  |||     ||| |
Db     522 KDELEQRMSALQESRR-ELMVQLEGLMKLLKEEELKQGTQGASSPRSSPSHTISRPIPMP 580
```

```
Qy     277 NAKLENSALLTVEPSPQDKNKGF-------FVDESEPLLRCD-------STSSGSSALSR 322
Db     581 IRSA--SACPTPTHTPQDSLTGVGGDVQEAFAQSSRRNLRSDLLVAADSITNTMSSLVKE 638

Qy     323 NGSFITKEKKDTVLRQVRLDPCDLQPIFDDM 353
Db     639 LNSEVASETESTV------DSEFSRPQFEDL 663


RESULT   4
POLG_PPVRA
ID   POLG_PPVRA     STANDARD;      PRT;  3140 AA.
AC   P17767;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER
DE   COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN
DE   1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2);
DE   GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
DE   (EC 3.4.22.44) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION
DE   PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48);
DE   COAT PROTEIN (CP)].
OS   Plum pox potyvirus (strain Rankovic) (PPV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC   Potyvirus.
OX   NCBI_TaxID=12214;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89370814; PubMed=2773595;
RA   Lain S., Riechmann J.L., Garcia J.A.;
RT   "The complete nucleotide sequence of plum pox potyvirus RNA.";
RL   Virus Res. 13:157-172(1989).
RN   [2]
RP   SEQUENCE OF 1778-2342 FROM N.A.
RX   MEDLINE=89268456; PubMed=2658302;
RA   Garcia J.A., Riechmann J.L., Lain S.;
RT   "Proteolytic activity of the plum pox potyvirus NIa-like protein in
RT   Escherichia coli.";
RL   Virology 170:362-369(1989).
RN   [3]
RP   SEQUENCE OF 2263-3140 FROM N.A.
RA   Lain S., Riechmann J.L., Mendez E., Garcia J.A.;
RT   "Nucleotide sequence of the 3' terminal region of plum pox potyvirus
RT   RNA.";
RL   Virus Res. 10:325-342(1988).
CC   -!- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC       MAY BE INVOLVED IN REPLICATION.
CC   -!- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: NUCLEAR INCLUSION PROTEIN A HYDROLYSES
CC       GLUTAMINYL BONDS, AND ACTIVITY IS FURTHER RESTRICTED BY
CC       PREFERENCES FOR THE AMINO ACIDS IN P6 -P1' THAT VARY WITH THE
CC       SPECIES OF POTYVIRUS, E.G. GLU-XAA-XAA-TYR-XAA-GLN+(SER OR GLY)
CC       FOR THE ENZYME FROM TOBACCO ETCH VIRUS. THE NATURAL SUBSTRATE IS
CC       THE VIRAL POLYPROTEIN, BUT OTHER PROTEINS AND OLIGOPEPTIDES
CC       CONTAINING THE APPROPRIATE CONSENSUS SEQUENCE ARE ALSO CLEAVED.
CC   -!- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC   -!- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC       POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC       PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC       INDIVIDUAL PROTEINS.
CC   -!- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC   -!- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC   -!- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21847; AAA85458.1; -.
DR   EMBL; M26965; AAA47085.1; -.
DR   PIR; A60009; GNVSRA.
DR   MEROPS; C04.001; -.
DR   MEROPS; C06.001; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001730; Peptidase_C4.
DR   InterPro; IPR001456; Peptidase_C6.
DR   InterPro; IPR002540; Poty_P1.
DR   InterPro; IPR001592; Poty_coat.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00863; Peptidase_C4; 1.
DR   Pfam; PF00851; Peptidase_C6; 1.
DR   Pfam; PF00767; Poty_coat; 1.
DR   Pfam; PF01577; Poty_P1; 1.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   PRINTS; PR00966; NIAPOTYPTASE.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
KW   Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW   Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW   ATP-binding.
FT   CHAIN       1    301       N-TERMINAL PROTEIN.
FT   CHAIN     302    914       HELPER COMPONENT PROTEINASE.
FT   CHAIN     915      ?       PROTEIN P3.
FT   CHAIN       ?   1168       6 KDA PROTEIN 1.
FT   CHAIN    1169   1803       CYTOPLASMIC INCLUSION PROTEIN.
FT   CHAIN    1804   1856       6 KDA PROTEIN 2.
FT   CHAIN    1857      ?       GENOME-LINKED PROTEIN.
FT   CHAIN       ?   2292       NUCLEAR INCLUSION PROTEIN A.
FT   CHAIN    2293   2810       NUCLEAR INCLUSION PROTEIN B.
FT   CHAIN    2811   3140       COAT PROTEIN.
FT   BINDING  1919   1919       COVALENT LINKAGE OF VIRAL RNA (BY
FT                              SIMILARITY).
FT   NP_BIND  1253   1260       ATP (POTENTIAL).
SQ   SEQUENCE   3140 AA;  355577 MW;  5F3DBB07982CF3F5 CRC64;


  Query Match             4.7%;  Score 98.5;  DB 1;  Length 3140;
  Best Local Similarity   20.4%;  Pred. No. 48;
  Matches   64;  Conservative   52;  Mismatches  103;  Indels   95;  Gaps   15;

Qy      92 KHFDINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKK--GPRQDP----------SAIV 139
Db    1024 KMFD-------MLIVFSLLLSIGATCNSMINEHKHLKQLAADREDKKRFKRLQVLHTRLS 1076

Qy     140 EKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFS 199
Db    1077 EKVG----CTPT--ADEFLEYVGGENPDLLKHAEDLIGDG-----QVVVHQSKRD----- 1120

Qy     200 NGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND----VVEKIRGLMEDTTQ-- 253
Db    1121 --SQANLERVVA-------------FVALVMMLFDSERSDGVYKILNKLKGIMGSVDQAV 1165

Qy     254 -----------LETDKLALP--MSPSPLSPS-PIPSPNAKLENSALLTVEPSPQDKNKGF 299
Db    1166 QHQSLDDIEDILDEKKLTVDFVLQSNEVAPTVPFDSTFEKWWTNQLETGNVIPHYRTEGH 1225

Qy     300 FVD---------------ESEPLLRCDSTSSGSSA-----LSRNGSFI----TKEKKDT 334
Db    1226 FLEFTRENAAHIANEVMHGSHQDILIRGAVGSGKSTGLPFHLSKKGHVLLIEPTRPLAEN 1285

Qy     335 VLRQVRLDPCDLQP 348
Db    1286 VCKQLRGQPFNVNP 1299


RESULT   5
APG1_YEAST
```

```
ID   APG1_YEAST     STANDARD;      PRT;   897 AA.
AC   P53104;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   AUTOPHAGY SERINE/THREONINE-PROTEIN KINASE APG1 (EC 2.7.1.-).
GN   APG1 OR AUT3 OR YGL180W OR G1615.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RX   MEDLINE=97197971; PubMed=9046087;
RA   Coglievina M., Klima R., Bertani I., Delneri D., Zaccaria P.,
RA   Bruschi C.V.;
RT   "Sequencing of a 40.5 kb fragment located on the left arm of
RT   chromosome VII from Saccharomyces cerevisiae.";
RL   Yeast 13:55-64(1997).
RN   [2]
RP   SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS.
RC   STRAIN=X2180;
RX   MEDLINE=97368130; PubMed=9224897;
RA   Matsuura A., Tsukada M., Wada Y., Ohsumi Y.;
RT   "Apg1p, a novel protein kinase required for the autophagic process in
RT   Saccharomyces cerevisiae.";
RL   Gene 192:245-250(1997).
CC   -!- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN THE REGULATION OF THE
CC       AUTOPHAGIC PROCESS, A PROTEIN DEGRADATION PROCESS. FOUND TO BE
CC       ESSENTIAL FOR THE FORMATION OF AUTOPHAGOSOMES. REQUIRED FOR THE
CC       MAINTENANCE OF CELL VIABILITY UNDER STARVATION.
CC   -!- PTM: AUTOPHOSPHORYLATED. MAY BE REGULATED BY A SECOND PROTEIN
CC       KINASE. THE PHOSPHORYLATION STATE MAY PLAY A ROLE IN THE INDUCTION
CC       OF PROTEIN DEGRADATION UPON STARVATION.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       APG1/UNC-51/ULK1 SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X91489; CAA62794.1; -.
DR   EMBL; Z72702; CAA96892.1; -.
DR   EMBL; D29991; BAA21481.1; -.
DR   HSSP; P24941; 1AQ1.
DR   SGD; S0003148; APG1.
DR   InterPro; IPR000719; Euk_pkinase.
DR   InterPro; IPR002290; Ser_thr_kin_actsite.
DR   Pfam; PF00069; pkinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Phosphorylation.
FT   DOMAIN       24    325       PROTEIN KINASE.
FT   NP_BIND      30     38       ATP (BY SIMILARITY).
FT   BINDING      54     54       ATP (BY SIMILARITY).
FT   ACT_SITE    172    172       BY SIMILARITY.
FT   MUTAGEN     211    211       D->A: FAILS TO COMPLEMENT APG1 MUTANT.
FT   MUTAGEN     237    237       E->R: FAILS TO COMPLEMENT APG1 MUTANT.
SQ   SEQUENCE   897 AA;  101717 MW;  7F4C785AA3A7CC46 CRC64;

  Query Match             4.7%;  Score 98;  DB 1;  Length 897;
  Best Local Similarity   24.0%;  Pred. No. 9.6;
  Matches   49;  Conservative   32;  Mismatches   79;  Indels   44;  Gaps    8;

Qy     228 LISALRQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLT 287
           |:|     |   ||||          :| : ||  :: :  :|:||  | :    : ||
Db     419 LVSDRSFEREYVVVEK--------KSVEVNSLADEVAQAGFNPNPIKHPTSTQNQNVLLN 470

Qy     288 VEPSPQDKNKGFFVDESE--PLLRCDSTSSGSSALSRNGSFITKEKKDTVL--------- 336
           : ||   |: :| :: |   ||| |:||| |  ||  | : :    : |
Db     471 EQFSP--NNQQYFQNQGENPRLLRATSSSSGGSDGSRRPSLVDRRLSISSLNPSNALSRA 528

Qy     337 ------------------RQVRLDPCDLQPIFDDML-HFLNPEELRVIEEIPQAED-KLD 376
                            :|:   |  | : :  | | |    :  |: :   |  |||
Db     529 LGIASTRLFGGANQQQQQQQITSSPPYSQTLLNSQLFHELTENIILRIDHLQHPETLKLD 588

Qy     377 RLFEIIGVKSQEASQTLLDSVYSH 400
               |: :    |:  :  |||:
Db     589 NT-NIVSILESLAAKAFV--VYSY 609


RESULT   6
MC3A_MOUSE
ID   MC3A_MOUSE     STANDARD;      PRT;  1971 AA.
AC   Q9WUU9;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   80 KDA MCM3-ASSOCIATED PROTEIN (GANP PROTEIN).
GN   MCM3AP OR GANP OR MAP80.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C X NZB;
RX   MEDLINE=20197882; PubMed=10733502;
RA   Kuwahara K., Yoshida M., Kondo E., Sakata A., Watanabe Y., Abe E.,
RA   Kouno Y., Tomiyasu S., Fujimura S., Tokuhisa T., Kimura H., Ezaki T.,
RA   Sakaguchi N.;
RT   "A novel nuclear phosphoprotein, GANP, is up-regulated in centrocytes
RT   of the germinal center and associated with MCM3, a protein essential
RT   for DNA replication.";
RL   Blood 95:2321-2328(2000).
CC   -!- FUNCTION: MAY BE INVOLVED IN THE NUCLEAR LOCALIZATION PATHWAY OF
CC       MCM3 (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH MCM3.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC.
CC   -!- PTM: PHOSPHORYLATED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE SAC3 FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AJ006590; CAB44241.1; -.
DR   MGD; MGI:1930089; Mcm3ap.
KW   Nuclear protein; Phosphorylation.
SQ   SEQUENCE   1971 AA;  217138 MW;  5F342E256C007E24 CRC64;

  Query Match             4.6%;  Score 97;  DB 1;  Length 1971;
  Best Local Similarity   21.8%;  Pred. No. 33;
  Matches   67;  Conservative   45;  Mismatches   98;  Indels   98;  Gaps   17;

Qy     146 KSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTAD 205
           ||:  :: ||:|||     |:  |:| | |     |:|  :|   ::| :         |
Db     405 KSVRQSKRREEWIYSLG--GVSSLELTAIQC----KNIPDYL---NDRAILEKHFSKIAK 455

Qy     206 HERAYA------ALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDTT------Q 253
           :| :       |: |:      |:||:             :| :|| :|       :
Db     456 VQRVFTRRSKKLAVIHF-FDHASAALAR-------------KKGKGLHKDVVIFWHKKK 500
```

```
Qy    254 LETDKLALPM----------SPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
          :   |   |:                ||   ||:  |  :    :||: : ||  |
Db    501 ISPSKKLFPLKEKLGESEASQGIEDSPFQHSPLSKPIVRPAAGSLLS-KSSPVKKPSLLK 559

Qy    301 VD--ESEPLLRCDSTSSGSSALSRNGSFITK-----------------EKKDTVLRQV 339
          :   |::|    || | ||  |   || ::                   :::| ::||
Db    560 MHQFEADPF---DSGSEGSEGL---GSCVSSLSTLIGTVADTSEEKYRLLDQRDRIMRQA 613

Qy    340 RLDPCDLQPIFDDMLHFLN------PE----------ELRVIEEIPQAEDKLDRLFEIIG 383
          |:   ||    |    |:       ||          :| | | :|   |::|    :
Db    614 RVKRTDL----DKARAFVGTCPDMCPEKERYLRETRSQLSVFEVVP-GTDQVDHAAAVKE 668

Qy    384 VKSQEASQ 391
            | |
Db    669 YSRSSADQ 676


RESULT   7
YEAC_SCHPO
ID   YEAC_SCHPO     STANDARD;      PRT;   969 AA.
AC   O14077;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PUTATIVE ZINC-PROTEASE C2E11.12C (EC 3.4.99.-).
GN   SPAC2E11.12C OR SPACUNK4.12C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC       INSULINASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL031180; CAA20142.1; -.
DR   InterPro; IPR001431; Peptidase_M16.
DR   Pfam; PF00675; Peptidase_M16; 1.
DR   PROSITE; PS00143; INSULINASE; 1.
KW   Hypothetical protein; Hydrolase; Metalloprotease; Zinc.
FT   METAL        68     68       ZINC (BY SIMILARITY).
FT   ACT_SITE     71     71       BY SIMILARITY.
FT   METAL        72     72       ZINC (BY SIMILARITY).
FT   METAL       149    149       ZINC (BY SIMILARITY).
SQ   SEQUENCE   969 AA;  112144 MW;  283A70BFBCFF3913 CRC64;


  Query Match             4.6%;  Score 96;  DB 1;  Length 969;
  Best Local Similarity   19.0%;  Pred. No. 15;
  Matches   77;  Conservative   60;  Mismatches  145;  Indels  124;  Gaps   14;

Qy     14 PKVLSSIQEGTVPDNTSSARGKEDVNKTLPN----------LQVVNHQQGPHHRHILKLL 63
          || :  : |   |:|  :      : | | |          |: :: |     ||
Db    422 PKGIQEVVESLRPNNFFAILAAHSIEKGLDNKEKFYGIDYGLEDLDSQFIDSLLHI---- 477

Qy     64 PSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRKS 123
                 |:|: :  |              || :|| : :    |   : | :: ::
Db    478 ---------KTSSELYLPL------------ANEFIPWSLEVEKQPVTTKLKVPNLVRN 515

Qy    124 SRTLKKGPRQDPSAIVEK---------------------------------------- 141
```

```
            : ::   ::| :  | |
Db    516 DKFVRLWHKKDDTFWVPKANVFINFISPIARRSPKVSVSTTLYTRLIEDALGEYSYPASL 575

Qy    142 AGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDI------YQFLCNASEREV 195
          |||  |::|:    : |  |     | | ::  :| :  :|:       :: | |  |:|:
Db    576 AGLSFSLSPS---TRGIILCISGFTDKLHVLLEKVVAMMRDLKVHPQRFEILKNRLEQEL 632

Qy    196 AAFSNGYTADHERAYAALQH---WTIRGPEASLAQLISALRQHRRNDVVEKIRGLM---- 248
            :      |   ||    |   |       | |:|  |::  :  |: : |  |:
Db    633 KDY------DALEAYHRSNHVLTWLSEPHSWSNAELREAIKDVQVGDMSDFISDLLKQNF 686

Qy    249 ----------EDTTQL--ETDKLALPMSPSPLSPSPIPSPNAKL----ENSALLTVEPS 291
                     ||   |    ||   : | |:  | :    | :     |    || |:
Db    687 LESLVHGNYTEEDAKNLIESAQKL---IDPKPVFASQLSRKRAIIVPEGGNYIYKTVVPN 743

Qy    292 PQDKNKGFFVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLR 337
          ::||     :   : : |   ||  ||:|    | ||   ::||
Db    744 KEEKNSAIMYNLQ--ISQLDDERSG--ALTRLARQIMKEPTFSILR 785


RESULT   8
RRP6_YEAST
ID   RRP6_YEAST     STANDARD;      PRT;   733 AA.
AC   Q12149;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   EXOSOME COMPLEX EXONUCLEASE RRP6 (EC 3.1.13.-) (RIBOSOMAL RNA
DE   PROCESSING PROTEIN 6).
GN   RRP6 OR UNC733 OR YOR001W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pettersson B., Sterky F., Uhlen M.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sterky F., Uhlen M.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=98250784; PubMed=9582370;
RA   Briggs M.W., Burkard K.T.D., Butler J.S.;
RT   "Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen,
RT   is essential for efficient 5.8 S rRNA 3' end formation.";
RL   J. Biol. Chem. 273:13255-13263(1998).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=99396719; PubMed=10465791;
RA   Allmang C., Petfalski E., Podtelejnikov A., Mann M., Tollervey D.,
RA   Mitchell P.;
RT   "The yeast exosome and human PM-Scl are related complexes of 3'-->5'
RT   exonucleases.";
RL   Genes Dev. 13:2148-2158(1999).
CC   -!- FUNCTION: COMPONENT OF THE NUCLEAR EXOSOME 3->5 EXORIBONUCLEASE
CC       COMPLEX. REQUIRED FOR THE 3'PROCESSING OF THE 7S PRE-RNA TO THE
CC       MATURE 5.8S RRNA. HAS A 3'-5' EXONUCLEASE ACTIVITY.
CC   -!- SUBUNIT: COMPONENT OF THE EXOSOME MULTIENZYME RIBONUCLEASE COMPLEX
CC       COMPOSED OF AT LEAST 11 PROTEINS: RRP4, RRP40, RRP41/SKI6, RRP42,
CC       RRP43, RRP44/DIS3, RRP45, RRP46, MTR3, CSL4 AND RRP6 (ONLY IN THE
CC       NUCLEAR COMPLEX). ALSO ASSOCIATED WITH THE GTPASE RAN (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS (POTENTIAL).
CC   -!- SIMILARITY: CONTAINS 1 HRDC DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; Z74909; CAA99189.1; -.
DR   EMBL; U43491; AAC49480.1; -.
DR   SGD; S0005527; RRP6.
DR   InterPro; IPR002121; HRDC.
DR   InterPro; IPR002562; 3_5_exonuclease.
DR   Pfam; PF01612; 3_5_exonuclease; 1.
DR   Pfam; PF00570; HRDC; 1.
DR   SMART; SM00474; 35EXOc; 1.
DR   SMART; SM00341; HRDC; 1.
KW   Hydrolase; Nuclease; Exonuclease; rRNA processing; Nuclear protein;
KW   RNA-binding.
FT   DOMAIN      435    515       HRDC.
SQ   SEQUENCE   733 AA;  84038 MW;  B19A0B2ED74C6DA7 CRC64;


  Query Match             4.6%;  Score 95.5;  DB 1;  Length 733;
  Best Local Similarity   19.7%;  Pred. No. 11;
  Matches   77;  Conservative   54;  Mismatches  144;  Indels  115;  Gaps   14;

Qy      77 PIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPS 136
           |:  |    : | : |   :| : |   ::    |  |:    :: |      |     ||
Db     353 PLSKPMTAYARADTHFLLNIYDQLRNKLIESNKLAGVLYESRNVAKRRFEYSKYRPLTPS 412

Qy     137 AIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASER--- 193
           : |        :| :   |     : |   : |  :   ||:|:     : | |
Db     413 SEV--------YSPIEKESPWKILMYQYNIPPEREVLVRELYQWRDLIARRDDESPRFVM 464

Qy     194 -----------------EVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHR 236
                            | : :|| | :| |  | |           || ||
Db     465 PNQLLAALVAYTPTDVIGVVSLTNGVT-EHVRQNAKL-----------LANLI------- 505

Qy     237 RNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVE-PSPQDK 295
             | :  |:   |: |                    ||||   | :   | |:  |  :|
Db     506 -RDALRNIKNTNEEAT------------------PIPSSETKADGILLETISVPQIRDV 545

Qy     296 NKGFFVDESEPLLRCDSTSSGSSALSRNGSFITKEK-----------KDTVLRQVRLDP 343
           : | |         |:|  | | |     | |   |             || :  :|: :
Db     546 MERFSV-------LCNSNISKSRAKPVTNSSILLGKILPREEHDIAYSKDGLPNKVKTED 598

Qy     344 CDLQP-------------IFDDMLHFLNP---EELRVIE--EIPQAEDKLDRLFEIIGVK 385
             ::              ||:     : |   ||:: ::    |    ::| | ::: :|
Db     599 IRIRAQNFKSALANLEDIIFEIEKPLVVPVKLEEIKTVDPASAPNHSPEIDNLDDLVVLK 658

Qy     386 ---------SQEASQTLLDSV-YSHLPDLL 405
                   ::|   |  |:| || :|::|
Db     659 KKNIQKKQPAKEKGVTEKDAVDYSKIPNIL 688


RESULT   9
MK21_YEAST
ID   MK21_YEAST     STANDARD;      PRT;  1025 AA.
AC   Q12176;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   RIBOSOME BIOGENESIS PROTEIN MAK21.
GN   MAK21 OR YDR060W OR D4237 OR YD9609.14.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96381250; PubMed=8789263;
RA   Brandt P., Ramlow S., Otto B., Bloecker H.;
RT   "Nucleotide sequence analysis of a 32,500 bp region of the right arm
RT   of Saccharomyces cerevisiae chromosome IV.";
RL   Yeast 12:85-90(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=99003241; PubMed=9786894;
RA   Edskes H.K., Ohtake Y., Wickner R.B.;
RT   "Mak21p of Saccharomyces cerevisiae, a homolog of human CAATT-binding
RT   protein, is essential for 60 S ribosomal subunit biogenesis.";
RL   J. Biol. Chem. 273:28912-28920(1998).
CC   -!- FUNCTION: REQUIRED FOR 60S RIBOSOMAL SUBUNIT SYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE CBF / MAK21 FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; X84162; CAA58976.1; -.
DR   EMBL; Z74356; CAA98878.1; -.
DR   EMBL; Z49209; CAA89089.1; -.
DR   SGD; S0002467; MAK21.
KW   Ribosome biogenesis; Nuclear protein.
FT   DOMAIN       74     77       POLY-ASP.
FT   DOMAIN      117    122       POLY-ASP.
FT   DOMAIN      979    986       POLY-GLU.
SQ   SEQUENCE   1025 AA;  116676 MW;  FB80378727ED71D8 CRC64;


  Query Match             4.6%;  Score 95.5;  DB 1;  Length 1025;
  Best Local Similarity   20.2%;  Pred. No. 18;
  Matches   96;  Conservative   59;  Mismatches  172;  Indels  149;  Gaps   20;

Qy       4 TESNSS----ASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQ-------Q 52
             :|:| :    :|:| |: | :::         : :|| || |   | : ||
Db       2 SENNGNPLDLSSLRNKISSKLRDNNSKKAKKTHKGK-DV-KASSNSKKVNEDIRREALAL 59

Qy      53 GPHHRHILKLLPSMEATGGEKSSTPI----------KGPKRGHPRQNLHKHFDINEHLPW 102
           |      |||:  :       ||                || |  :  ||  |:   ::|
Db      60 GASEED-LKLIQGLSDDDDAKSEQEFDAVADEDADDKGFK--NDLQNFMKNVGFDQH--- 113

Qy     103 MIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCN 162
                                 |     |     |    |:|  | | :|
Db     114 ------------------------KLEDVDDDDIEEESTSSKESKIPAQEKE------- 141

Qy     163 GHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSN-GYTADHERAYAALQHWTIRGP 221
           |    :      :  ||      : : | ||:|         ::| |              |
Db     142 -HAQSNIASSTIEKTSQ-----ESIDNGSEQEENTVEEANLSSDQE------------P 182

Qy     222 EASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPS----PLSPSPIPSPN 277
           |:  |:          :    ||  ||:  || : :||| :|        || |    : :
Db     183 ESESAE---------KEKKEEKDGGLITQTTIISSDKLIIPYDKPWYEIPLDPQVGQNDD 233

Qy     278 A---------KLENSALLTVEPSPQDKNKGFFVDESE--------------------PLL 308
                     ||      |:|   |   : |  | |:                     ||
Db     234 VEELSKEQIEKLFERGKQTLEADNQTYYEEFTKDSSQAKFMSQILSDGTLNDKISAVTLL 293

Qy     309 RCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNP-EELRVIEE 367
             ||    : :|    |:  |: ::: |:        |  : |  |: | |  :||  :
Db     294 IQDSPLHNTKSLETLVSYCGKKSRNSALQ-------SLNALKDLFLNGLLPNRKLRYFKN 346

Qy     368 IP--------------QAEDKLDRLF----EIIGVKSQEASQTLLDSVYSHLPDLL 405
           |                || | :||    |:: | | :    :   : :|: |||
Db     347 QPGLSMMLNKKTLAIFYFEDYLKKLFFRVLEVLEVLSHDPIIHVRLQILNHVFDLL 402


RESULT  10
CYP1_YEAST
ID   CYP1_YEAST     STANDARD;      PRT;  1483 AA.
```

```
AC   P12351;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CYP1 ACTIVATORY PROTEIN.
GN   CYP1 OR HAP1 OR YLR256W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89125585; PubMed=2851658;
RA   Creusot F., Verdiere J., Gaisne M., Slonimski P.P.;
RT   "CYP1 (HAP1) regulator of oxygen-dependent gene expression in yeast.
RT   I. Overall organization of the protein sequence displays several
RT   novel structural domains.";
RL   J. Mol. Biol. 204:263-276(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89106221; PubMed=2643482;
RA   Pfeifer K., Kim K.-S., Kogan S., Guarente L.;
RT   "Functional dissection and sequence of yeast HAP1 activator.";
RL   Cell 56:291-301(1989).
RN   [3]
RP   MUTANT CYP1-18.
RX   MEDLINE=89125586; PubMed=2851659;
RA   Verdiere J., Gaisne M., Guiard B., Defranoux N., Slonimski P.P.;
RT   "CYP1 (HAP1) regulator of oxygen-dependent gene expression in yeast.
RT   II. Missense mutation suggests alternative Zn fingers as
RT   discriminating agents of gene control.";
RL   J. Mol. Biol. 204:277-282(1988).
RN   [4]
RP   DNA-BINDING.
RX   MEDLINE=97042358; PubMed=8887558;
RA   Zhang L., Guarente L.;
RT   "The C6 zinc cluster dictates asymmetric binding by HAP1.";
RL   EMBO J. 15:4676-4681(1996).
RN   [5]
RP   STRUCTURE BY NMR OF 60-100.
RX   MEDLINE=96275662; PubMed=8683583;
RA   Timmerman J., Vuidepot A.-L., Bontems F., Lallemand J.-Y., Gervais M.,
RA   Shechter E., Guiard B.;
RT   "1H, 15N resonance assignment and three-dimensional structure of CYP1
RT   (HAP1) DNA-binding domain.";
RL   J. Mol. Biol. 259:792-804(1996).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 55-135 OF MUTANT HAP1-18.
RX   MEDLINE=99101378; PubMed=9886287;
RA   King D.A., Zhang L., Guarente L., Marmorstein R.;
RT   "Structure of HAP1-18-DNA implicates direct allosteric effect of
RT   protein-DNA interactions on transcriptional activation.";
RL   Nat. Struct. Biol. 6:22-27(1999).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 56-135 COMPLEXED WITH DNA.
RX   MEDLINE=99101385; PubMed=9886294;
RA   King D.A., Zhang L., Guarente L., Marmorstein R.;
RT   "Structure of a HAP1-DNA complex reveals dramatically asymmetric DNA
RT   binding by a homodimeric protein.";
RL   Nat. Struct. Biol. 6:64-71(1999).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 55-135 OF MUTANT HAP1-PC7.
RX   MEDLINE=20482201; PubMed=11024163;
RA   Lukens A.K., King D.A., Marmorstein R.;
RT   "Structure of HAP1-PC7 bound to DNA: implications for DNA recognition
RT   and allosteric effects of DNA-binding on transcriptional activation.";
RL   Nucleic Acids Res. 28:3853-3863(2000).
CC   -!- FUNCTION: REGULATION OF OXYGEN DEPENDENT GENE EXPRESSION. IT
CC       MODULATES THE EXPRESSION OF ISO-1 AND ISO-2 CYTOCHROME C. BINDS
CC       TO THE SEQUENCE 5'-CGGNNNTNNCGG-3'.
CC   -!- SUBUNIT: BINDS DNA AS A HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- MISCELLANEOUS: CYP1-18 MUTANT ACTIVATE THE EXPRESSION OF CYP3
CC       (ISO-2) WHILE REDUCING THAT OF CYC1 (ISO-1).
CC   -!- MISCELLANEOUS: HEME IS AN EFFECTOR MOLECULE FOR CYP1/HAP1. THE
CC       REPEAT REGION (SEE FT TABLE) MEDIATES HEME INDUCTION BY MASKING
CC       THE DNA-BINDING DOMAIN IN THE ABSENCE OF INDUCER.
CC   -!- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC       CLUSTER DOMAIN.
CC   -!- SIMILARITY: CONTAINS 7 HEME REGULATORY MOTIFS (HRM).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X13793; CAA32032.1; -.
DR   EMBL; J03152; AAA34662.1; -.
DR   PIR; S05804; RGBYH1.
DR   PDB; 1PYC; 01-AUG-96.
DR   PDB; 2HAP; 18-MAY-99.
DR   PDB; 1HWT; 18-MAY-99.
DR   PDB; 1QP9; 09-OCT-00.
DR   TRANSFAC; T00346; -.
DR   SGD; S0004246; HAP1.
DR   InterPro; IPR001138; Zn2_CY6_fungal.
DR   Pfam; PF00172; Zn_clus; 1.
DR   SMART; SM00066; GAL4; 1.
DR   PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR   PROSITE; PS50048; ZN2_CY6_FUNGAL_2; 1.
KW   Transcription regulation; Activator; DNA-binding; Nuclear protein;
KW   Zinc; Metal-binding; Repeat; Heme; 3D-structure.
FT   DNA_BIND      64     93       ZN(2)-CYS(6), FUNGAL-TYPE.
FT   METAL         64     64       ZINC 1 AND 2.
FT   METAL         67     67       ZINC 1.
FT   METAL         74     74       ZINC 1.
FT   METAL         81     81       ZINC 1 AND 2.
FT   METAL         84     84       ZINC 2.
FT   METAL         93     93       ZINC 2.
FT   DOMAIN       177    189       POLY-GLN.
FT   DOMAIN       280    285       HRM 1 (POTENTIAL).
FT   DOMAIN       299    304       HRM 2 (POTENTIAL).
FT   DOMAIN       323    328       HRM 3 (POTENTIAL).
FT   DOMAIN       347    352       HRM 4 (POTENTIAL).
FT   DOMAIN       389    394       HRM 5 (POTENTIAL).
FT   DOMAIN       415    420       HRM 6 (POTENTIAL).
FT   DOMAIN      1192   1197       HRM 7 (POTENTIAL).
FT   VARIANT       63     63       S -> R (IN CYP1-18 MUTANT).
FT   CONFLICT     145    145       T -> I (IN REF. 2).
FT   CONFLICT     323    323       K -> R (IN REF. 2).
FT   CONFLICT     455    455       S -> N (IN REF. 2).
FT   CONFLICT     508    508       V -> M (IN REF. 2).
FT   CONFLICT     587    587       N -> K (IN REF. 2).
FT   CONFLICT     883    883       D -> N (IN REF. 2).
FT   CONFLICT     960    960       H -> S (IN REF. 2).
FT   CONFLICT    1151   1151       D -> N (IN REF. 2).
FT   CONFLICT    1157   1157       S -> P (IN REF. 2).
FT   CONFLICT    1305   1305       N -> Y (IN REF. 2).
SQ   SEQUENCE   1483 AA;  164150 MW;  570ECA20E4ED71C6 CRC64;


  Query Match             4.6%;  Score 95.5;  DB 1;  Length 1483;
  Best Local Similarity   20.8%;  Pred. No. 29;
  Matches   80;  Conservative   62;  Mismatches  149;  Indels   93;  Gaps   19;

Qy      11 SVRPKVLSSIQEGTVPDNTSSARGKEDVNKT----LPNL-QVVNHQQGPHHRHILKLLPS 65
           :|   | | | :| :|   :|  |  | |      || |  ::|| | |::::      :
Db     880 NVNDVVSSLINKGLLP---TSEGGSVDSNNDEIYGLPKLPDILNH--GQHNQNLY----- 929

Qy      66 MEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRKSSR 125
             | |   ||: |   |  | :: :   ::|:  ::|:||  ::          :
Db     930 --ADGRNTSSSDI-DKKLDLPHESTTRALFFSKHMTIRMLLYLLNYILF---------TH 977
```

```
Qy     126 TLKKGPRQDPSAIVEKAGLKKSMTPTQN--REKWIYYCNGHGID--------ILKLVAAQ 175
Db     978 YEPMGSEDPGTNILAKEYAQEALNFAMDGYRNCMIFFNNIRNTNSLFDYMNVILSYPCLD 1037

Qy     176 VGSQWKDIYQFL--------CN--ASEREVAAFSNGYTADHERAYAALQHWTIRGPEASL 225
Db    1038 IGHR---SLQFIVCLILRAKCGPLTGMRESSIITNG--------------TSSGFNSSV 1079

Qy     226 AQLISALRQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSAL 285
Db    1080 EDEDVKVKQESSDEL--KKDDFMKDVNLDSGDSLAEILMSRMLLFQKLTKQLSKKYNYAI 1137

Qy     286 LTVEPSPQDKNKGFFVDESE-PLLRCDSTSSGSS---------ALSRNGSFITKEKKDTV 335
Db    1138 ------RMNKSTGFFVSLLDTPSKKSDSKSGGSSFMLGNWKHPKVSNMSGFLAGDKD--- 1188

Qy     336 LRQVRLDPCDLQPIFDDMLHFLNP 359
Db    1189 ----QLQKC---PVYQDALGFVSP 1205
```

RESULT 11
POLG_PPVSK
ID POLG_PPVSK STANDARD; PRT; 3140 AA.
AC Q84934;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER
DE COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN
DE 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2);
DE GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
DE (EC 3.4.22.44) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION
DE PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48);
DE COAT PROTEIN (CP)].
OS Plum pox potyvirus (strain SK 68) (PPV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=103927;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167908; PubMed=8122394;
RA Palkovics L., Burgyan J., Balazs E.;
RT "Comparative sequence analysis of four complete primary structures of
RT plum pox virus strains.";
RL Virus Genes 7:339-347(1993).
CC -!- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -!- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC MAY BE INVOLVED IN REPLICATION.
CC -!- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -!- CATALYTIC ACTIVITY: NUCLEAR INCLUSION PROTEIN A HYDROLYSES
CC GLUTAMINYL BONDS, AND ACTIVITY IS FURTHER RESTRICTED BY
CC PREFERENCES FOR THE AMINO ACIDS IN P6 -P1' THAT VARY WITH THE
CC SPECIES OF POTYVIRUS, E.G. GLU-XAA-XAA-TYR-XAA-GLN+(SER OR GLY)
CC FOR THE ENZYME FROM TOBACCO ETCH VIRUS. THE NATURAL SUBSTRATE IS
CC THE VIRAL POLYPROTEIN, BUT OTHER PROTEINS AND OLIGOPEPTIDES
CC CONTAINING THE APPROPRIATE CONSENSUS SEQUENCE ARE ALSO CLEAVED.
CC -!- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -!- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC INDIVIDUAL PROTEINS.
CC -!- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -!- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -!- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC ---

CC ---
DR EMBL; M92280; AAB05823.1; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR001730; Peptidase_C4.
DR InterPro; IPR001456; Peptidase_C6.
DR InterPro; IPR002540; Poty_P1.
DR InterPro; IPR001592; Poty_coat.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00863; Peptidase_C4; 1.
DR Pfam; PF00851; Peptidase_C6; 1.
DR Pfam; PF00767; Poty_coat; 1.
DR Pfam; PF01577; Poty_P1; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR PRINTS; PR00966; NIAPOTYPTASE.
DR SMART; SM00487; DEXDc; 1.
DR SMART; SM00490; HELICc; 1.
KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW ATP-binding.
FT CHAIN 1 301 N-TERMINAL PROTEIN.
FT CHAIN 302 914 HELPER COMPONENT PROTEINASE.
FT CHAIN 915 ? PROTEIN P3.
FT CHAIN ? 1168 6 KDA PROTEIN 1.
FT CHAIN 1169 1803 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1804 1856 6 KDA PROTEIN 2.
FT CHAIN 1857 ? GENOME-LINKED PROTEIN.
FT CHAIN ? 2292 NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2293 2810 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2811 3140 COAT PROTEIN.
FT BINDING 1919 1919 COVALENT LINKAGE OF VIRAL RNA (BY
FT SIMILARITY).
FT NP_BIND 1253 1260 ATP (POTENTIAL).
SQ SEQUENCE 3140 AA; 355992 MW; 837A5A692B56436A CRC64;

Query Match 4.6%; Score 95.5; DB 1; Length 3140;
Best Local Similarity 20.1%; Pred. No. 81;
Matches 63; Conservative 54; Mismatches 102; Indels 95; Gaps 15;

```
Qy      92 KHFDINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKK--GPRQDP----------SAIV 139
Db    1024 KMFD-------MLIIFGLLLSIGATCNSMINEHKHLKQVAADREDKKRFKRLQVLYTRLL 1076

Qy     140 EKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFS 199
Db    1077 EKIG----CTPT--ADEFLEYVQGENPDLSKYAEDLIGDG-----QVVVHQSKRD----- 1120

Qy     200 NGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND----VVEKIRGLMEDTTQ-- 253
Db    1121 --SQANLERV------------AAFVALVMMLFDSERSDGVYKILNKLKGVMGSIDQTV 1165

Qy     254 ----------LETDKLALP--MSPSPLSPS-PIPSPNAKLENSALLTVEPSPQDKNKGF 299
Db    1166 HHQNLDDIEDMLDEKKLTVDFVLQSNEVAPTVPFDSTFEKWWTNQLETGNVIPHYRTEGH 1225

Qy     300 FVD---------------ESEPLLRCDSTSSGSSA-----LSRNGSFI----TKEKKDT 334
Db    1226 FLEFTRENAAHIANEVMHGSHQDILIRGAVGSGKSTGLPFHLSKKGHVLLIEPTRPLAEN 1285

Qy     335 VLRQVRLDPCDLQP 348
Db    1286 VCKQLRGQPFNVNP 1299
```

RESULT 12
YJ9C_YEAST
ID YJ9C_YEAST STANDARD; PRT; 707 AA.
AC P47166;
DT 01-FEB-1996 (Rel. 33, Created)

```
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   HYPOTHETICAL 81.2 KDA PROTEIN IN NMD5-HOM6 INTERGENIC REGION.
GN   YJR134C OR J2120.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rose M., Koetter P., Entian K.D.;
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: SOME, TO PARAMYOSINS AND MYOSINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z49634; CAA89665.1; -.
DR   EMBL; Z49635; CAA89667.1; -.
DR   SGD; S0003895; YJR134C.
KW   Hypothetical protein; Coiled coil.
FT   DOMAIN      122    473       COILED COIL (POTENTIAL).
FT   DOMAIN      594    706       COILED COIL (POTENTIAL).
SQ   SEQUENCE   707 AA;  81171 MW;  BA47D1A7265BF3D9 CRC64;


  Query Match             4.5%;  Score 95;  DB 1;  Length 707;
  Best Local Similarity   21.6%;  Pred. No. 12;
  Matches   52;  Conservative   44;  Mismatches   85;  Indels   60;  Gaps   11;

Qy     223 ASLAQLISALRQHRRND---VVEKI----------RGLMED--TTQLETDKLALPMSP- 265
           : | |  : |: |: |    :::||           | :||  |   |   |:| ||
Db     454 SQLEQKENELKPHQENSNEKIIDKIPVELTDSLNSMEGNIEDEWTLPQENSMLSLSMSKL 513

Qy     266 SPLSPSPIPSPNAKLENSALLTVEPSPQD-KNKGFFVD----------ESEPLLRCD 311
             |   |   |     :  : : |    | ||  | :|            | | :   :
Db     514 GELESDPSLKPIYNESHETICSEESQHFDRKNVDFSIDDIPEEAAALQAIREGESMNSLN 573

Qy     312 STS----SGSSALSRNGSFITK---EKKDTVLRQVRLDPCDLQPIFDDMLH---FLNPEE 361
           :||     |  || :   |:     |  | |:::  :    : ::|:::|      | |
Db     574 NTSIPYRRASVQLSNSNGHISAHLVNKLSTELKRLEGELSASKELYDNLLKEKTKANDEI 633

Qy     362 LRVIEE----------------IPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDL 404
           ||::||                : | : ||: . :::| |:::   :  |::  | | ::
Db     634 LRLLEENDKFNEVNKQKDDLLKRVEQMQSKLETSLQLLGEKTEQVEE--LENDVSDLKEM 691

Qy     405 L 405
           :
Db     692 M 692


RESULT  13
POLG_PPVD
ID   POLG_PPVD      STANDARD;      PRT;  3141 AA.
AC   P13529; Q84929; P89038;
DT   01-JAN-1990 (Rel. 13, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER
DE   COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN
DE   1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2);
DE   GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
DE   (EC 3.4.22.44) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION
DE   PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48);
DE   COAT PROTEIN (CP)].
OS   Plum pox potyvirus (strain D) (PPV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC   Potyvirus.
OX   NCBI_TaxID=12212;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90098790; PubMed=2602121;
RA   Teycheney P.Y., Tavert G., Delbos R., Ravelonandro M., Dunez J.;
RT   "The complete nucleotide sequence of plum pox virus RNA (strain D).";
RL   Nucleic Acids Res. 17:10115-10116(1989).
RN   [2]
RP   SEQUENCE OF 2810-3141 FROM N.A., AND SEQUENCE OF 2812-2828.
RA   Ravelonandro M., Varveri C., Delbos R., Dunez J.;
RT   "Nucleotide sequence of the capsid protein gene of plum pox
RT   potyvirus.";
RL   J. Gen. Virol. 69:1509-1516(1988).
RN   [3]
RP   REVISIONS TO C-TERMINUS.
RA   le Gall O.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
CC       TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC   -!- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC       MAY BE INVOLVED IN REPLICATION.
CC   -!- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: NUCLEAR INCLUSION PROTEIN A HYDROLYSES
CC       GLUTAMINYL BONDS, AND ACTIVITY IS FURTHER RESTRICTED BY
CC       PREFERENCES FOR THE AMINO ACIDS IN P6 -P1' THAT VARY WITH THE
CC       SPECIES OF POTYVIRUS, E.G. GLU-XAA-XAA-TYR-XAA-GLN+(SER OR GLY)
CC       FOR THE ENZYME FROM TOBACCO ETCH VIRUS. THE NATURAL SUBSTRATE IS
CC       THE VIRAL POLYPROTEIN, BUT OTHER PROTEINS AND OLIGOPEPTIDES
CC       CONTAINING THE APPROPRIATE CONSENSUS SEQUENCE ARE ALSO CLEAVED.
CC   -!- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC   -!- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC       POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC       PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC       INDIVIDUAL PROTEINS.
CC   -!- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC   -!- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC   -!- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16415; CAA34437.1; -.
DR   EMBL; D00298; BAA00210.1; -.
DR   PIR; S06929; GNVSPD.
DR   PIR; JA0078; JA0078.
DR   MEROPS; C04.001; -.
DR   MEROPS; C06.001; -.
DR   MEROPS; S30.001; -.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001730; Peptidase_C4.
DR   InterPro; IPR001456; Peptidase_C6.
DR   InterPro; IPR002540; Poty_P1.
DR   InterPro; IPR001592; Poty_coat.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00863; Peptidase_C4; 1.
DR   Pfam; PF00851; Peptidase_C6; 1.
DR   Pfam; PF00767; Poty_coat; 1.
DR   Pfam; PF01577; Poty_P1; 1.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   PRINTS; PR00966; NIAPOTYPTASE.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
KW   Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW   Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW   ATP-binding.
```

```
FT   CHAIN         1    301       N-TERMINAL PROTEIN.
FT   CHAIN       302    915       HELPER COMPONENT PROTEINASE.
FT   CHAIN       916      ?       6 KDA PROTEIN 1.
FT   CHAIN         ?   1169       PROTEIN P3.
FT   CHAIN      1170   1804       CYTOPLASMIC INCLUSION PROTEIN.
FT   CHAIN      1805   1857       6 KDA PROTEIN 2.
FT   CHAIN      1858      ?       GENOME-LINKED PROTEIN.
FT   CHAIN         ?   2293       NUCLEAR INCLUSION PROTEIN A.
FT   CHAIN      2294   2811       NUCLEAR INCLUSION PROTEIN B.
FT   CHAIN      2812   3141       COAT PROTEIN.
FT   BINDING    1920   1920       COVALENT LINKAGE OF VIRAL RNA (BY
FT                                SIMILARITY).
FT   NP_BIND    1254   1261       ATP (POTENTIAL).
SQ   SEQUENCE   3141 AA;  355569 MW;  6C3641C404414DBB CRC64;


  Query Match             4.5%;  Score 95;  DB 1;  Length 3141;
  Best Local Similarity   20.1%;  Pred. No. 88;
  Matches   61;  Conservative   52;  Mismatches  102;  Indels   88;  Gaps   14;

Qy     103 MIVLFLLLVLVVIVVCSIRKSSRTLKK--GPRQDP----------SAIVEKAGLKKSMTP 150
           |:::| ||: :     |:    : ||:    |:|           : : || |     ||
Db    1029 MLIVFGLLLSIGATCNSMINEHKHLKQLAADREDKKRFKRLQVLHTRLSEKVG----CTP 1084

Qy     151 TQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAY 210
           |   :::: |  |   |:||     :|       | : : |:|:          |: ||
Db    1085 T--ADEFLEYVGGENPDLLKHAEDLIGDG-----QVVVHQSKRD-------SQANLERVV 1130

Qy     211 AALQHWTIRGPEASLAQLISALRQHRRND----VVEKIRGLMEDTTQ------------- 253
           |                |:  |    |:|    :: |::|:|    |
Db    1131 A-------------FVALVMMLFDSERSDGVYKILNKLKGIMGSVDQAVHHQSLDDIEDI 1177

Qy     254 LETDKLALP--MSPSPLSPS-PIPSPNAKLENSALLTVEPSPQDKNKGFFVD-------- 302
           |:  || :   :  : ::|: |  |   |-   : | |    |  : :| |::
Db    1178 LDEKKLTVDFVLQSNEVAPTVPFDSTFEKWWTNQLETGNVIPHYRTEGHFLEFTRENAAH 1237

Qy     303 ---------ESEPLLRCDSTSSGSSA-----LSRNGSFI----TKEKKDTVLRQVRLDPCD 345
                   : :|    :  || |      ||: |  :     |:   : | :|:|  | :
Db    1238 IANEVMHGSHQDILIRGAVGSGKSTGLPFHLSKKGHVLLIEPTRPLAENVCKQLRGQPFN 1297

Qy     346 LQP 348
           : |
Db    1298 VNP 1300


RESULT  14
YCX9_CHLVU
ID   YCX9_CHLVU     STANDARD;      PRT;   249 AA.
AC   O20173;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL 27.9 KDA PROTEIN IN TRNR-MINE INTERGENIC REGION (ORF249).
OS   Chlorella vulgaris.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC   Chlorellaceae; Chlorella.
OX   NCBI_TaxID=3077;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IAM C-27 / TAMIYA;
RX   MEDLINE=97303241; PubMed=9159184;
RA   Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA   Tsudzuki J., Nakashima K.; Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA   Inamura A., Yoshinaga K., Sugiura M.;
RT   "Complete nucleotide sequence of the chloroplast genome from the green
RT   alga Chlorella vulgaris: the existence of genes possibly involved in
RT   chloroplast division.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB001684; BAA57948.1; -.
KW   Hypothetical protein; Chloroplast.
SQ   SEQUENCE   249 AA;  27944 MW;  8A89631F981E4AE8 CRC64;


  Query Match             4.5%;  Score 94;  DB 1;  Length 249;
  Best Local Similarity   26.4%;  Pred. No. 3.4;
  Matches   46;  Conservative   21;  Mismatches   51;  Indels   56;  Gaps    8;

Qy     257 DKLALPMSP----SPLSPSP-----IPS---PNAKLENSALLTVEPSPQDK--------- 295
           ||  |   |    ||    |     :||   ||| ||:    |: |  | :
Db      34 DKSFLDPHPVARVSPFGSRPPFLLMVPSTSVPNASLESGQSSTISPRRQKESSIPVLQAS 93

Qy     296 -----NKGFFVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIF 350
                |  |   ||  | :  |:||      | |  ||:::  :     |: | :| |
Db      94 GSVSANKTFQSTESSALHQPKSSSS-----ERTGKIITRQRPHSGPTSPRVTPGELLPTL 148

Qy     351 DDMLHFLNPEELRVIEEIPQAEDKLDRLFEIIGVKSQEA-------SQTLLDSV 397
                  | |:: | :  |: ||       |:::       || |||::
Db     149 ---------EALKLQEALDFAQLKL---------KTRQVRYTINLRSQKLLDTI 184


RESULT  15
CLC1_MOUSE
ID   CLC1_MOUSE     STANDARD;      PRT;   994 AA.
AC   Q64347;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CHLORIDE CHANNEL PROTEIN, SKELETAL MUSCLE (CHLORIDE CHANNEL PROTEIN
DE   1) (CLC-1).
GN   CLCN1 OR CLC1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=97468267; PubMed=9321463;
RA   Schnuelle V., Antropova O., Gronemeier M., Wedemeyer N., Jockusch H.,
RA   Bartsch J.W.;
RT   "The mouse Clc1/myotonia gene: ETn insertion, a variable AATC repeat,
RT   and PCR diagnosis of alleles.";
RL   Mamm. Genome 8:718-725(1997).
RN   [2]
RP   SEQUENCE OF 350-467 FROM N.A.
RX   MEDLINE=92065955; PubMed=1659665;
RA   Steinmeyer K., Klocke R., Ortland C., Gronemeier M., Jockusch H.,
RA   Gruender S., Jentsch T.J.;
RT   "Inactivation of muscle chloride channel by transposon insertion in
RT   myotonic mice.";
RL   Nature 354:304-308(1991).
CC   -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
CC       SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
CC       MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
CC       TRANSEPITHELIAL TRANSPORT.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN SKELETAL MUSCLES.
CC   -!- SIMILARITY: TO OTHER CHLORIDE CHANNELS.
CC   -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z95127; CAB08359.1; -.
DR   EMBL; Z95128; CAB08359.1; JOINED.
DR   EMBL; Z95129; CAB08359.1; JOINED.
DR   EMBL; Z95130; CAB08359.1; JOINED.
DR   EMBL; Z95131; CAB08359.1; JOINED.
DR   EMBL; Z95132; CAB08359.1; JOINED.
DR   EMBL; Z95133; CAB08359.1; JOINED.
DR   EMBL; Z95134; CAB08359.1; JOINED.
DR   EMBL; Z95135; CAB08359.1; JOINED.
DR   EMBL; Z95136; CAB08359.1; JOINED.
DR   EMBL; Z95137; CAB08359.1; JOINED.
DR   EMBL; Z95138; CAB08359.1; JOINED.
DR   EMBL; Z95139; CAB08359.1; JOINED.
DR   EMBL; Z95140; CAB08359.1; JOINED.
DR   EMBL; Z95141; CAB08359.1; JOINED.
DR   EMBL; Z95142; CAB08359.1; JOINED.
DR   EMBL; Z95143; CAB08359.1; JOINED.
DR   EMBL; Z95144; CAB08359.1; JOINED.
DR   EMBL; Z95145; CAB08359.1; JOINED.
DR   EMBL; Z95146; CAB08359.1; JOINED.
DR   EMBL; Z95147; CAB08359.1; JOINED.
DR   EMBL; Z95148; CAB08359.1; JOINED.
DR   EMBL; Z95149; CAB08359.1; JOINED.
DR   EMBL; X62897; CAA44686.1; -.
DR   EMBL; X62895; CAA44684.1; -.
DR   EMBL; X62896; CAA44685.1; -.
DR   MGD; MGI:88417; Clcn1.
DR   InterPro; IPR000644; CBS.
DR   InterPro; IPR001807; Volt_Cl_channel.
DR   Pfam; PF00571; CBS; 2.
DR   Pfam; PF00654; voltage_CLC; 1.
DR   PRINTS; PR00762; CLCHANNEL.
DR   PRINTS; PR01112; CLCHANNEL1.
DR   SMART; SM00116; CBS; 1.
KW   Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW   CBS domain.
FT   DOMAIN        1    117       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    118    137       1 (POTENTIAL).
FT   TRANSMEM    159    182       2 (POTENTIAL).
FT   TRANSMEM    206    227       3 (POTENTIAL).
FT   TRANSMEM    235    254       4 (POTENTIAL).
FT   TRANSMEM    266    290       5 (POTENTIAL).
FT   TRANSMEM    305    323       6 (POTENTIAL).
FT   TRANSMEM    348    368       7 (POTENTIAL).
FT   TRANSMEM    391    414       8 (POTENTIAL).
FT   TRANSMEM    456    475       9 (POTENTIAL).
FT   TRANSMEM    478    496       10 (POTENTIAL).
FT   TRANSMEM    524    545       11 (POTENTIAL).
FT   TRANSMEM    553    572       12 (POTENTIAL).
FT   DOMAIN      573    844       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    845    863       13 (POTENTIAL).
FT   DOMAIN      864    994       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      607    662       CBS 1.
FT   DOMAIN      825    877       CBS 2.
SQ   SEQUENCE   994 AA;  109993 MW;  2136BC3E4131D58D CRC64;


  Query Match             4.5%;  Score 94;  DB 1;  Length 994;
  Best Local Similarity   23.4%;  Pred. No. 22;
  Matches   40;  Conservative   24;  Mismatches   63;  Indels   44;  Gaps    6;

Qy     249 EDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLL 308
           ||  :  : |: ||::|:|  ||| | |:          : ||  ::  |      :|
Db     710 EDEDEDLSRKMELPLTPAPPPPSPPPPPSQ-------FPIAPSNPEEPNGPLPSHKQPPE 762

Qy     309 RCDSTSSGSSALSR-----NGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELR 363
             ||    ||   |      |   :|:|| |         |   : |:|    :|||:
Db     763 ASDSADQRSSTFQRLLHCLLGKAHSKKKKITQ---------DSTDLVDNM----SPEEIE 809

Qy     364 VIEE-------------IPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHL 401
             |              | |:       |:::   :   : ||   :  ||
Db     810 AWEREQLSQPVCFDCCCIDQSP------FQLVEQTTLHKTHTLFSLLGLHL 854


Search completed: January  4, 2002, 08:39:30
Job time: 123 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: January 4, 2002, 08:36:52 ; Search time 25.5 Seconds
(without alignments)
2323.149 Million cell updates/sec

Title: US-09-733-956-2
Perfect score: 2091
Sequence: 1 MNSTESNSSASVRPKVLSSI.....SQEASQTLLDSVYSHLPDLL 405

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2091	100.0	655	4	O75509	O75509 homo sapien
2	1898	90.8	655	11	Q9EPU5	Q9epu5 mus musculu
3	1537	73.5	630	13	Q98SM6	Q98sm6 gallus gall
4	120.5	5.8	1059	5	Q9VVC6	Q9vvc6 drosophila
5	119	5.7	2781	4	Q9UIG2	Q9uig2 homo sapien
6	117	5.6	1151	5	Q9VAY4	Q9vay4 drosophila
7	116	5.5	599	5	Q9U3W0	Q9u3w0 drosophila
8	115.5	5.5	2183	11	O88783	O88783 mus musculu
9	115	5.5	927	10	Q9LF24	Q9lf24 arabidopsis
10	114.5	5.5	290	4	Q9Y3Z8	Q9y3z8 homo sapien
11	113.5	5.4	412	4	Q9H5E0	Q9h5e0 homo sapien
12	113	5.4	599	5	Q9V680	Q9v680 drosophila
13	112	5.4	1822	11	O35412	O35412 rattus norv
14	111	5.3	1138	4	O43166	O43166 homo sapien
15	111	5.3	1783	4	O95321	O95321 homo sapien
16	111	5.3	1804	4	Q9UNU4	Q9unu4 homo sapien
17	108.5	5.2	1180	4	Q92625	Q92625 homo sapien
18	107	5.1	309	12	Q89402	Q89402 paramecium
19	107	5.1	461	5	Q21004	Q21004 caenorhabdi
20	107	5.1	500	10	Q9LVQ7	Q9lvq7 arabidopsis
21	106.5	5.1	425	3	Q9C0Q7	Q9c0q7 neurospora
22	105.5	5.0	452	10	Q9SCQ6	Q9scq6 arabidopsis
23	105.5	5.0	524	2	Q9ZKU3	Q9zku3 helicobacte
24	105.5	5.0	1405	10	O50052	O50052 arabidopsis
25	104.5	5.0	607	3	O94395	O94395 schizosacch
26	104.5	5.0	3140	12	Q84925	Q84925 plum pox po
27	103	4.9	519	10	Q9FYW4	Q9fyw4 lycopersico
28	103	4.9	849	5	O77068	O77068 trichomonas
29	103	4.9	931	5	O96524	O96524 trichomonas
30	103	4.9	1231	11	O55092	O55092 cavia porce
31	103	4.9	1583	4	O15045	O15045 homo sapien
32	103	4.9	2168	5	Q9VQM0	Q9vqm0 drosophila
33	103	4.9	2176	5	O46112	O46112 drosophila
34	102.5	4.9	1385	5	Q9W4Z0	Q9w4z0 drosophila
35	101.5	4.9	494	5	Q23654	Q23654 caenorhabdi
36	101.5	4.9	514	5	Q10956	Q10956 caenorhabdi
37	101.5	4.9	527	2	O25564	O25564 helicobacte
38	101.5	4.9	739	5	Q9VZ86	Q9vz86 drosophila
39	101.5	4.9	919	4	Q9NQS7	Q9nqs7 homo sapien
40	101	4.8	16215	5	Q9NFS3	Q9nfs3 drosophila
41	100	4.8	498	2	Q9I4Y4	Q9i4y4 pseudomonas
42	99.5	4.8	872	3	Q9USS2	Q9uss2 schizosacch
43	99.5	4.8	887	5	Q9N627	Q9n627 leishmania
44	99	4.7	604	3	O74491	O74491 schizosacch
45	99	4.7	895	2	Q9CLK7	Q9clk7 pasteurella

ALIGNMENTS

RESULT 1
O75509
ID O75509 PRELIMINARY; PRT; 655 AA.
AC O75509;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TNFR-RELATED DEATH RECEPTOR-6 (DJ181J13.1) (DR6 (TNFR-RELATED DEATH
DE RECEPTOR-6)).
GN DR6 OR DJ181J13.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pan G., Bauer J.H., Haridas V., Wang S., Liu D., Ni J., Yu G.,
RA Vincenz C., Aggarwal B.B., Dixit V.M.;
RT "Identification and functional characterization of DR6, a novel death
RT domain-containing TNF receptor.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Parker A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF068868; AAC34583.1; -.
DR EMBL; AL096801; CAB75692.1; -.
DR HSSP; P07174; 1NGR.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00020; TNFR_c6; 4.
DR ProDom; PD000771; TNFR_c6; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00208; TNFR; 4.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 655 AA; 71844 MW; 48939391C4852A33 CRC64;

```
  Query Match           100.0%;  Score 2091;  DB 4;  Length 655;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-166;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 310

Qy      61 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSI 370

Qy     121 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 RKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQW 430

Qy     181 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 KDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDV 490

Qy     241 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 VEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFF 550

Qy     301 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 VDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPE 610

Qy     361 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     611 ELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 655


RESULT   2
Q9EPU5
ID   Q9EPU5      PRELIMINARY;      PRT;   655 AA.
AC   Q9EPU5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DR6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=KIDNEY;
RA   Isogai D., Ichino M., Yoshinari M., Yamaura A., Kurokawa F.,
RA   Minami M.;
RT   "Mouse DR6: mouse homolog of human TNFR-related death receptor-6
RT   (DR6).";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF322069; AAG38116.1; -.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001368; TNFR_c6.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00020; TNFR_c6; 4.
DR   ProDom; PD000771; TNFR_c6; 1.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00208; TNFR; 4.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR   PROSITE; PS50050; TNFR_NGFR_2; 1.
SQ   SEQUENCE   655 AA;  71909 MW;  1A97C1432799E4FB CRC64;


  Query Match           90.8%;  Score 1898;  DB 11;  Length 655;
  Best Local Similarity  92.4%;  Pred. No. 5.4e-150;
  Matches  375;  Conservative    7;  Mismatches   22;  Indels    2;  Gaps    2;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           ||||:|||:|||| || | |:|||||||||||  |||  |:||||   | ||| |||||||
Db     251 MNSTDSNSTASVRTKVPSGIEEGTVPDNTSSTSGKEGTNRTLPNPPQVTHQQAPHHRHIL 310

Qy      61 KLLP-SMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCS 119
           |||| ||||| |||||| || |||||||| |||||||||| ||||||||||:|||||
Db     311 KLLPSSMEAT-GEKSSTAIKAPKRGHPRQNAHKHFDINEHLPLMIVLFLLLVLVLIVVCS 369

Qy     120 IRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQ 179
           |||||||||||||||||||||||||||:||||||||||| ||||||||||||||||||
Db     370 IRKSSRTLKKGPRQDPSAIVEKAGLKKSLTPTQNREKWIYYRNGHGIDILKLVAAQVGSQ 429

Qy     180 WKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND 239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 WKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRND 489

Qy     240 VVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGF 299
           |||||||||||||||||||||||||||||||:|||| ||||| |||||||| |||| |
Db     490 VVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPMPSPNVKLENSTLLTVEPSPLDKNKCF 549

Qy     300 FVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNP 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     550 FVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHILNP 609

Qy     360 EELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 405
           |||||||||||||||||||||||||||||||||||||||||||||
Db     610 EELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL 655


RESULT   3
Q98SM6
ID   Q98SM6      PRELIMINARY;      PRT;   630 AA.
AC   Q98SM6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DEATH RECEPTOR 6 (FRAGMENT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bridgham J.T., Johnson A.L.;
RT   "Expression of DR6 in the ovary.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF349908; AAK29666.1; -.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   630 AA;  68977 MW;  53C53BD1C4B25567 CRC64;


  Query Match           73.5%;  Score 1537;  DB 13;  Length 630;
  Best Local Similarity  73.7%;  Pred. No. 6.8e-120;
  Matches  314;  Conservative   25;  Mismatches   61;  Indels   26;  Gaps    5;

Qy       1 MNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL 60
           :||:  : |:|  |:| :   | ||  | :|| |      :| :       |   |:|
Db     210 LNSSVFDLSSSPAPRVSNGTAEPTVDYNDTSANGTVGAPGSLSSAGTAGQAQSYRHKHTS 269

Qy      61 KLL---PSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVV 117
           : :   |: |  |||||| | : |:|| |. |:|:|||||||||||||||||||||||||
Db     270 QAMGKQPAQEMAGGEKSSIPYRPPRRGPP--NVHQHFDINEHLPWMIVLFLLLVLVVIVV 327

Qy     118 CSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVG 177
           ||:|||||| ||| |||||||||| :||| |||||||||||||||||||||||| ||||
Db     328 CSVRKSSRTXKKGXRQDPSAIVEKAIMKKSTTPTQNREKWIYYCNGHGIDILKLVXAQVG 387

Qy     178 SQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRR 237
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     388 SQWKDIYQFLCNASEREVAAFSNGYAADHERAYAALQHWTIRGPEASLAQLISALRQHRR 447

Qy     238 NDVVEKIRGLMEDTT-----------------QLETDKLALPMSPSPLSPSPIPSPNAK 279
           |||||||||||||||                  :|| ||||||:||||||| | |||  |
```

```
Db    448 NDVVEKIRGLMEDTTPVQMQPQWQTQDCSNDDGKLEGDKLALPVSPSPLSPVPTPSP--K 505

Qy    280 LENSALLTVEPSPQDKNKGFFVDESEPLLRCDSTSSGSSALSRNGSFITKEKKDTVLRQV 339
          ::|:||||||| :| | ||||| ||||||||||||||||||| ||||||||||||||||
Db    506 PPDAAVLTVEPSPSEK-KCFFVDEXEPLLRCDSTSSGSSALSRTGSFITKEKKDTVLRQV 564

Qy    340 RLDPCDLQPIFDDMLHFLNPEELRVIEEIPQAEDKLDRLFEIIGVKSQEASQTLLDSVYS 399
          |||||||||||||||| |||||| |||||| |||||| |||| ||||||||||||||||
Db    565 RLDPCDLQPIFDDMLHILNPEELHVIEEIPAAEDKLDXLFEIAGVKSQEASQTLLDSVYS 624

Qy    400 HLPDLL 405
          ||||||
Db    625 HLPDLL 630


RESULT   4
Q9VVC6
ID   Q9VVC6      PRELIMINARY;      PRT;  1059 AA.
AC   Q9VVC6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG14060 PROTEIN.
GN   CG14060.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003525; AAF49387.1; -.
DR   FlyBase; FBgn0036683; CG14060.

SQ   SEQUENCE   1059 AA;  115746 MW;  F8E12EF89F38A18B CRC64;


  Query Match             5.8%;  Score 120.5;  DB 5;  Length 1059;
  Best Local Similarity   22.2%;  Pred. No. 0.21;
  Matches   88;  Conservative   47;  Mismatches  136;  Indels  125;  Gaps   18;

Qy      2 NSTESNSSASVRPKVLSSIQEGTVPDNTSS--ARGKEDVNKTLPNLQVVNHQQGPHHRHI 59
            ::: | : | :| :|  || | :: |:||      |||    ||   :   :     |
Db    570 SASPSPTKAQIREEV--SIDEESIKDSTSEDMVTRKEDFTLDLPKEVIPRAEDLEDSSHA 627

Qy     60 L----------KLLPSMEATGGE-KSSTPIKG----------------------PKRGH 85
          |          |:| ::||   : |   |:                        |
Db    628 LLVSPCFEADAKILDALEAVDLKLKVLNPLNDVTTADGERELLLAANDDFEQRILPNSSF 687

Qy     86 PRQNLHKHFDINEHL--------------PWMIVLFLLLVLVVIVVCSIRKSSRTLK-- 128
            |  :| :  |: |                 ||:    |: |  :  : |   : ||
Db    688 VYQAANKQYSNNQILLNQAVNVTLVNAIEGEQMIMEQKQLMKVRDLELTERDIDQPLKRA 747

Qy    129 -KGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWK------ 181
           | | :| :  :|   | : :||  :  |                    ||  | :|
Db    748 EKSPEKDDT--LEPESLYEPITPVMDETK--------------TEAQTSSLFKIYTKKS 790

Qy    182 ----DIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRR 237
              |||| :  ||  ||    | :   :|         :| | : :
Db    791 ELVEDIYQTVEEASTVEVK--PNNWLEGYE---------SIAGSQEACPLTYDGYESFAP 839

Qy    238 NDVVEKI----------RGLMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLT 287
           :||             |   ::  :|   |  |  || || |:| | | ||
Db    840 EEVVTTPTTPSATGTLGRNTRDELPELPKPKRVLNKSPIPLRPAP-PKPQAK-------- 890

Qy    288 VEPSPQDKN-----KGFFVDESEPLLRCDSTSSGSS 318
          | :|:|     || :  ||   : | :||| ||
Db    891 --ESEEDENIYDTIKGCY--ES---MHCKATSSSSS 919


RESULT   5
Q9UIG2
ID   Q9UIG2      PRELIMINARY;      PRT;  2781 AA.
AC   Q9UIG2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   BROMODOMAIN PHD FINGER TRANSCRIPTION FACTOR.
GN   BPTF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Jones M.H., Hamana N., Shimane M.;
RT   "Iedentification and charcterization BPTF a novel bromodomain
RT   transcription factor.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB032251; BAA89208.1; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001487; Bromodomain.
DR   InterPro; IPR001965; PHD.
DR   Pfam; PF00439; bromodomain; 1.
DR   Pfam; PF00628; PHD; 2.
DR   PRINTS; PR00503; BROMODOMAIN.
DR   PROSITE; PS00633; BROMODOMAIN_1; 1.
DR   PROSITE; PS50014; BROMODOMAIN_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   SMART; SM00297; BROMO; 1.
DR   SMART; SM00249; PHD; 2.
SQ   SEQUENCE   2781 AA;  311210 MW;  7B3BE0C1C4C3CFBE CRC64;


  Query Match             5.7%;  Score 119;  DB 4;  Length 2781;
  Best Local Similarity   18.0%;  Pred. No. 1;
  Matches   79;  Conservative   78;  Mismatches  169;  Indels  112;  Gaps   18;
```

```
Qy      15 KVLSSIQEGTVPDNTSSARG-KEDVNKTLPNLQVVNHQQGPHHRHILKLLPSMEATGGEK 73
            :||| ||    |      :    |   : :    || :|:    |        | :: : :::|    ::
Db    2304 QVLSQIQSQVVAQIQAQQSGVPQQIKLQLP-IQI--QQSSAVQTHQIQNVVTVQAASVQE 2360

Qy      74 SSTPIKGPKRGHPRQNLHKHFDI-NEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPR 132
             ::     |     ::     :  :|   ||                              |   ::|
Db    2361 QLQRVQ-QLRDQQQKKKQQQIEIKREH----------------TLQASNQSEIIQKQVV 2402

Qy     133 QDPSAIVEKAGLKKSMTPTQNRE-KWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNAS 191
             :|::|      |||||| :    |  :   :   ||         ::|  :    ::   :  |
Db    2403 MKHNAVIEHLKQKKSMTPAEREENQRMIVCN---QVMKYILDKIDKEEKQ--------- 2449

Qy     192 EREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDT 251
                              |   :    ::    :         |: |:    ::    | ::::| | |::
Db    2450 ----------AAKKRKREESVEQKRSKQNATKLSALLFKHKEQLRAEILKK-RALLDKD 2497

Qy     252 TQLET-------------DKLAL--------------PMSPSPLSPSPIPSPNAKLENS 283
             |:|                |  : |                  |: |:| :| |  |  |    ::::
Db    2498 LQIEVQEELKRDLKIKKEKDLMQLAQATAVAAPCPPVTPVLPAPPAPPPSPPPPPGVQHT 2557

Qy     284 ALLTVEPSPQDKNKGFFVDESEPLLRCDSTSSGSSALSRNGSFITKE-KKDTVLRQVRLD 342
             ||:      |      |         ||   |      :     |     :||  ||||  |
Db    2558 GLLSTPTLPVASQKRKREEKD---------SSSKSKKKKMISTTSKETKKDTKLY----- 2604

Qy     343 PCDLQPIFDDMLHFLN----------------PEELRVIEE--IPQAEDKLDRLFEI--I 382
             |   :   :|:    ::                       |  :|:|     ||   :     |   :   :   :
Db    2605 -CICKTPYDESKFYIGCDRCQNWYHGRCVGILQSEAELIDEYVCPQCQSTEDAMTVLTPL 2663

Qy     383 GVKSQEASQTLLDSVYSH 400
              |   |    :  :|  |:  :|
Db    2664 TEKDYEGLKRVLRSLQAH 2681
```

RESULT 6
Q9VAY4

```
ID   Q9VAY4      PRELIMINARY;      PRT;  1151 AA.
AC   Q9VAY4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG5514 PROTEIN.
GN   CG5514.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003763; AAF56763.1; -.
DR   FlyBase; FBgn0039560; CG5514.
DR   InterPro; IPR002965; P_rich_extensn.
DR   PRINTS; PR01217; PRICHEXTENSN.
SQ   SEQUENCE   1151 AA;  125450 MW;  AF8330378C58C1D1 CRC64;

  Query Match             5.6%;  Score 117;  DB 5;  Length 1151;
  Best Local Similarity   25.8%;  Pred. No. 0.47;
  Matches   54;  Conservative   29;  Mismatches   80;  Indels   46;  Gaps    9;

Qy     167 DILKLVAAQVGSQWK-DIYQFLCNASEREVAAFSNGYTADHERA---YAALQHWTIRGPE 222
            | :| : |:| ||   | ::| |  :: :                      | |    : | | ||   ||
Db       3 DFIKTLIAEVKSQGVFDEFRFNCCLADVDTKPAYQNVRTQVETAVNDFLAKQQWT---PE 59

Qy     223 ASLAQLISALRQH-RRNDVVEKIRGLMEDTTQLETDKLAL-----------------PM 263
             :  ||   ||:|    :||::| |: :   |:   |:|                        |
Db      60 TNKVQLRERLRKHLMDSDVLDK--GVDQIVDQVVNPKVATIFEPKIESIVYKYLGITPPA 117

Qy     264 SPSPLSPSPIPSPNAKLENSALLTVEP-----------SPQDKNKGFFVDESE-----P 306
             |: ||  |:|   |  |  |:|| ||                   || |  ::       ||      |
Db     118 RPTMLSAPPLPPFGAHLNGSSLLNVETTVGLLPTDLEQISP-DSDRATVKSESRDDELPP 176

Qy     307 LLRCDSTSSGSSALSRNGSFITKEKKDTV 335
              :  : ||     :|    :   ||: : |
Db     177 GVDDEDTSPSYELVSERKTLAIKEELNNV 205
```

RESULT 7
Q9U3W0

```
ID   Q9U3W0      PRELIMINARY;      PRT;   599 AA.
AC   Q9U3W0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SCAVENGER RECEPTOR-LIKE PROTEIN PRECURSOR.
GN   SR-CII OR CG8856.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pearson A.M., Trigatti B., Xu S., Krieger M.;
RT   "Identification of a family of genes related to the Drosophila
RT   scavenger receptor dSR-CI.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF216642; AAF19660.1; -.
DR   HSSP; P10998; 1VVD.
DR   FlyBase; FBgn0020377; Sr-CII.
DR   InterPro; IPR000998; MAM.
DR   InterPro; IPR001212; Somatomedin_B.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00629; MAM; 1.
DR   Pfam; PF01033; Somatomedin_B; 1.
```

```
DR   Pfam; PF00084; sushi; 2.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00137; MAM; 1.
DR   SMART; SM00201; SO; 1.
DR   PROSITE; PS50060; MAM_2; 1.
DR   PROSITE; PS00524; SOMATOMEDIN_B; UNKNOWN_1.
KW   Receptor.
FT   CHAIN        21    599       SCAVENGER RECEPTOR-LIKE PROTEIN.
SQ   SEQUENCE   599 AA;  67098 MW;  C211450347AA38F0 CRC64;


  Query Match             5.5%;  Score 116;  DB 5;  Length 599;
  Best Local Similarity   23.9%;  Pred. No. 0.24;
  Matches   55;  Conservative   38;  Mismatches   61;  Indels   76;  Gaps   12;

Qy       3 STESNSSASVRPKVLSSIQEG----TVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRH 58
           :| | |  :   || |   |     | |  :::       : | |  |:
Db     403 TTTSISRNTTTEKVFSISLEASRSFTTPRTSTN-------STTFPRTQISTK-------- 447

Qy      59 ILKLLPSMEATGGEKSSTPIKGPKRGHP----------------RQNLHK----HFDIN 97
             ||| : | |  | |: | :  |   |                  |:: |:      |:|
Db     448 --KLLNTHE-TSTETSTLPTRNKKLSQPSIFSSQHITVTTTTDRSRKDYHEDIAGQLDMN 504

Qy      98 EHLPW--MIVLFLLL-VLVVIVVCSIRKSSRTLKKGPRQDPSAIV--EKA--GLKKSMTP 150
            ::|   :||::||: |::|||| ::::  : | |   :|   :|  :||  ||:|
Db     505 TNIPHPALIVMYLLIGVILVIVVANLKQLCKILTKNSSRDNEKVVSFKKAFEGLRKPRRL 564

Qy     151 TQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSN 200
           :||       |||                : | ||:| : ::  | :
Db     565 SQN-------CNG------------------MDQHLCSAFDEDLDYFED 588


RESULT   8
O88783
ID   O88783      PRELIMINARY;      PRT;  2183 AA.
AC   O88783;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MURINE COAGULATION FACTOR V.
GN   F5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang T.L., Cui J., Rehumtulla A., Moussalli M., Kaufman R.J.,
RA   Ginsburg D.;
RT   "The structure and function of murine factor V and its inactivation by
RT   protein C.";
RL   Blood 91:0-0(0012).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Ginsburg D., Yang T.L., Cui J., Yang A.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U52925; AAC99553.1; -.
DR   HSSP; P00450; 1KCW.
DR   MGD; MGI:88382; F5.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 3.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
SQ   SEQUENCE   2183 AA;  247228 MW;  BF0A8AA723F60317 CRC64;


  Query Match             5.5%;  Score 115.5;  DB 11;  Length 2183;
  Best Local Similarity   19.9%;  Pred. No. 1.5;
  Matches   97;  Conservative   49;  Mismatches  197;  Indels  145;  Gaps   20;

Qy       5 ESNSSASVRPKVLSSIQEGTVPD--NTSSARGKEDVNKTLPNL----------QVVNHQQ 52
           :||||     ::|| |    : |   |    |       | ||           |:
Db     773 DSNSS-----RILSKIINNNLKDFQRTLPGSGATVAGTLLRNLIGLDENFVLNSSTEHRS 827

Qy      53 GPHHRH----------ILKLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHF-------D 95
             :| :           :: |||      | :     |  | | |    |:
Db     828 SSYHENDMENPQSNITMVYLLPLGPKGSGNREQDKPKTIKTGRPHMMKHRFSWMKAPAGK 887

Qy      96 INEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNRE 155
             |                   |  |:| :  |     || ::    ||: :|
Db     888 TGRH------------------SNPKNSYSGMKSEEDIPSELIP---LKQKITSKFLNR 925

Qy     156 KW----------IYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTAD 205
           :|          |   ||   |: ||                  |: : :     | :
Db     926 RWRVASEKGSYEIIAANGEDTDVDKLT---------------NSPQNQNITVPRGESTS 969

Qy     206 HERAYAALQHWTIRGPE-----ASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDKLA 260
           |          | | |      : :      :||   |   :| :  :    : :  |||
Db     970 HTN--------TTRKPSDLPTFSGVGHKSPHVRQEEENSGFQKRQLFIRTRKKKKNKKLA 1021

Qy     261 L--PMSPSPLSP----SPIPSPNAKLENSALL-------TVEPSPQDKNKGFFVDESEP- 306
           |  |:||    |     :  | |: :| | :||        : |     :     | | |
Db    1022 LHSPLSPRGFDPLRGHNHSPFPDRRLLNHSLLLHKSNETALSPDLNQTSPSMSTDRSLPD 1081

Qy     307 --LLRCDSTSSGSSALSRNGSFITKEKKDTVLRQ-----------VRLDPCDLQPIFD- 351
             : |   ||:|    |  :|   |  |                 |  | :|    |
Db    1082 YNQYSKNDTEQMSSSLDLYQSVPAEEHSPTFPAQDPDQTHSTTDPSYRSSPPELSQGLDY 1141

Qy     352 DMLHFLNPEELRVIEEIPQAEDKLD-------------RLFEIIGVKSQEASQTLLDSV 397
           |: |   |::: :    |    |                 || |    | |  ||:   :
Db    1142 DLSHDFYPDDIGLTSFFPDQSQKSSFSSDDDQAIPSSDLSLFTI----SPELDQTI---I 1194

Qy     398 YSHLPDLL 405
           |  |  ||
Db    1195 YPDLDQLL 1202


RESULT   9
Q9LF24
ID   Q9LF24      PRELIMINARY;      PRT;   927 AA.
AC   Q9LF24;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 104.1 KDA PROTEIN.
GN   T20K14_190.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA   Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL391143; CAC01757.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   927 AA;  104087 MW;  0BB628F06C516803 CRC64;


  Query Match             5.5%;  Score 115;  DB 10;  Length 927;
  Best Local Similarity   18.9%;  Pred. No. 0.51;
  Matches   83;  Conservative   51;  Mismatches  158;  Indels  146;  Gaps   15;

Qy      17 LSSIQEGTVPDN---TSSARGKEDVNKTLPNLQVVNHQQGPHHRHILKL---------- 62
           | |:  |    ||   |: :  |::  |:          |:|       :|
Db      45 LKSLPSGKASDNVGDTNISADKKETEKSKKKKTAKEKQRGVSSESSSRLSFSSSPCSSSF 104
```

```
Qy      63 --------LPSMEATGGEKSSTPIKGPKRGHPR-QNLHKHFDINEHLPWMIVLFLLLVLV 113
                   |  |      |:: |  | ||   |    || |
Db     105 SSADISTTASQFEQPGLSNGENPVREPTNGSPRWGGLMMPSDIRE--------------- 149

Qy     114 VIVVCSIRKSSRTL-KKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLV 172
           :|  || | :||  ::   | | :      | |  :|::|  :|       |  ::||
Db     150 -LVRSSIHKETRTRDEEALSQQPKSARANVSLLKESSPSRNSNEW-----SEGRRVVKL- 202

Qy     173 AAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISAL 232
                 ||  :|  :  |||
Db     203 --------KDSPRF--SYDERE----------------------------------- 214

Qy     233 RQHRRNDVVEKIRGLMEDTTQLETDKLALPMSPSPLSPSPIP-----------SPNAKL 280
                   |    :::| :|  |  :     :  | || |            |  |||
Db     215 --------TRKTGAKLKETPRLSLDSRSNSFRSARSSCSPEPQELVTGHRRTTSSVVAKL 266

Qy     281 ENSALLTVEPSPQDKNKGFFVDESEPLLRCDSTSSGSSALSRNGSFITKEKK-------- 332
               ::  ||      :  | |   |      ||    | |:  | : :|
Db     267 MGLEVIPDEPVTIQNRENRFCDSPRP------TSRVEVDLQRSRGFDSIKKMMPAKFPMK 320

Qy     333 -------DTVLRQVRLDPCDLQPIFDDM------LHFLNPE-ELRVIEEIPQAEDKLDRL 378
                |    ||::       :: ::      | |   | :|| :::| :| :|  :|
Db     321 ASPWAQVDGAKNQVKIPDATTLTVYGEIQKRLSQLEFKKSEKDLRALKQILEAMEKTQQL 380

Qy     379 FEIIGVKSQEASQTLLDS 396
                 |  : ::||  |
Db     381 IS----KDDDDNKTLCSS 394


RESULT  10
Q9Y3Z8
ID   Q9Y3Z8      PRELIMINARY;      PRT;   290 AA.
AC   Q9Y3Z8;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   HYPOTHETICAL 33.3 KDA PROTEIN (FRAGMENT).
GN   DKFZP564K112.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL049996; CAB43230.1; -.
KW   Hypothetical protein.
FT   NON_TER      1      1
SQ   SEQUENCE   290 AA;  33296 MW;  19298B9DDEDE1DAD CRC64;


  Query Match             5.5%;  Score 114.5;  DB 4;  Length 290;
  Best Local Similarity   21.1%;  Pred. No. 0.12;
  Matches   66;  Conservative   46;  Mismatches  104;  Indels   97;  Gaps   13;

Qy       2 NSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHILK 61
           :|:|   | : | ||  : ||:  : : : :::   |:       ::   ::
Db       6 SSSEDAESLAPRSKVKKIIQKDIIKEASEASKENRDI-------EISTEEE--------- 49

Qy      62 LLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLP--------WMIVLFLLLVLV 113
                : ||  | |: :|  || |- :::  :|    | :|        || :    | |
Db      50 -----KDTGDLKDSSLLK-TKRKHKKKHKERHKMGEEVIPLRVLSKSEWMDLKKEYLALQ 103

Qy     114 VIVVCSIRKSSRTLK-KGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYC------NGHGI 166
              : |::|:   :| :   :  | : :  |:|     | |||:    |       |  |
Db     104 KASMASLKKTISQIKSESEMETDSGVPQNTGMKNE--KTANREE----CRTQEKVNATGP 157

Qy     167 D-----ILKLVAAQVGSQWKDIYQFLCNASE-------------------REVAAFSNG 201
               |:|::: :     | :   |   ||                     :  |  |
Db     158 QFVSGVIVKIISTEPLPGRKQVRDTLAAISEVLYVDLLEGDTECHARFKTPEDAQAVINA 217

Qy     202 YT--------------ADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRGL 247
           ||               |||: |     |          | |   || : |   || ||
Db     218 YTEINKRHCWKLEILSGDHEQRY-----W----------QKILVDRQAKLNQPREKKRGT 262

Qy     248 MEDTTQLETDKLA 260
             :  || |  :||
Db     263 EKLITQAEKIRLA 275


RESULT  11
Q9H5E0
ID   Q9H5E0      PRELIMINARY;      PRT;   412 AA.
AC   Q9H5E0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CDNA: FLJ23531 FIS, CLONE LNG06065.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HUMAN LUNG;
RA   Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA   Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA   Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA   Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA   Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK027184; BAB15686.1; -.
SQ   SEQUENCE   412 AA;  46539 MW;  D72A6D830BB12B94 CRC64;


  Query Match             5.4%;  Score 113.5;  DB 4;  Length 412;
  Best Local Similarity   17.9%;  Pred. No. 0.23;
  Matches   70;  Conservative   59;  Mismatches  133;  Indels  129;  Gaps   12;

Qy       9 SASVRPKVL---SSIQEGTVPDNTSSARGKEDVNKTLPNLQVVN--HQQGPHHRHILKLL 63
           |   ||| |    |: :   | ::  :   :  ::   |:|  :    | |   |
Db      94 SLKFRPKQLFHPMSLLKHNPPTHSHPSPKLQSQSQPQSNVQGQSPVRVQSPSQTRIRPST 153

Qy      64 PSMEATGGE-------------KSSTPIKGPKRGHPR---QNLHKHFDINEHLPWMIVLF 107
           ||  : | :              :  | :: | :| |:   | : ||
Db     154 PSQLSPGQQSQVQTTTSQPIPIQPHTSLQIPSQGQPQSQPQVVMKH-------------- 199

Qy     108 LLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNRE-KWIYYCNGHGI 166
                                      :|:::|    ||||||  :  | : :  ||
Db     200 ----------------------------NAVIEHLKQKKSMTPAEREENQRMIVCN---- 227

Qy     167 DILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLA 226
             ::|.:  ::  : |                     |   :   :::    :    |:
Db     228 QVMKYILDKIDKEEKQ-------------------AAKKRKREESVEQKRSKQNATKLS 267

Qy     227 QLISALRQHRRNDVVEKIRGLMEDTTQLETDK---------------------------- 258
             |:   ::  | ::::| | |::   |:|  :
Db     268 ALLFKHKEQLRAEILKK-RALLDKDLQIEVQEELKRDLKIKKEKDLMQLAQATAVAAPCP 326

Qy     259 ---LALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLLRCDSTSS 315
               | |  |:| :| | | |   :::: ||:   |    |    :| :          ||
Db     327 PVTPAPPAPPAPPAPPPSPPPPPAVQHTGLLSTPTLPAASQKRKREEEKD--------SS 378

Qy     316 GSSALSRNGSFITKE-KKDTVLRQVRLDPCD 345
             |   :  |  :|| |||| |  :   | |
Db     379 SKSKKKKMISTTSKETKKDTKLYCICKTPYD 409


RESULT  12
Q9V680
ID   Q9V680      PRELIMINARY;      PRT;   599 AA.
AC   Q9V680;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG8856 PROTEIN.
GN   SR-CII OR CG8856.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003823; AAF58551.1; -.
DR   HSSP; P10998; 1VVD.
DR   FlyBase; FBgn0020377; Sr-CII.
DR   InterPro; IPR000998; MAM.
DR   InterPro; IPR001212; Somatomedin_B.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00629; MAM; 1.
DR   Pfam; PF01033; Somatomedin_B; 1.
DR   Pfam; PF00084; sushi; 2.
DR   SMART; SM00032; CCP; 2.
DR   SMART; SM00137; MAM; 1.
DR   SMART; SM00201; SO; 1.
DR   PROSITE; PS50060; MAM_2; 1.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 1.
SQ   SEQUENCE   599 AA;  67146 MW;  8A70752A7997EF89 CRC64;

  Query Match             5.4%;  Score 113;  DB 5;  Length 599;
  Best Local Similarity   23.9%;  Pred. No. 0.42;
  Matches   55;  Conservative   38;  Mismatches   61;  Indels   76;  Gaps   12;

Qy       3 STESNSSASVRPKVLSSIQEG----TVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRH 58
           :| | |  :   || |    |     | |  :::       : | |  |:
Db     403 TTTSISRNTTTEKVFSISLEASRSFTTPRTSTN-------STTFPRTQISTK-------- 447

Qy      59 ILKLLPSMEATGGEKSSTPIKGPKRGHP----------------RQNLHK----HFDIN 97
             ||| : | |  | |: | :  |   |                    |:: |:      |:|
Db     448 --KLLNTHE-TLTETSTIPTRNKKLSQPSIFSSQHITVTTTTDRSRKDYHEDIAGQLDMN 504

Qy      98 EHLPW--MIVLFLLL-VLVVIVVCSIRKSSRTLKKGPRQDPSAIV--EKA--GLKKSMTP 150
             ::|   :||::||: |::|||| ::::  : | |   :|   :|  :||  ||:|
Db     505 TNIPHPALIVMYLLIGVILVIVVANLKQLCKILSKNSSRDNEKVVSFKKAFEGLRKPRRL 564

Qy     151 TQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSN 200
           :||       |||                  : | ||:| : ::  | :
Db     565 SQN-------CNG------------------MDQHLCSAFDEDLDYFED 588


RESULT  13
O35412
ID   O35412      PRELIMINARY;      PRT;  1822 AA.
AC   O35412;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SPA-1 LIKE PROTEIN P1294.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takeuchi M., Ide N., Hata Y., Takai Y.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF026504; AAB81526.1; -.
DR   InterPro; IPR000331; Rap_GAP.
DR   InterPro; IPR001478; PDZ.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF02145; Rap_GAP; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   SMART; SM00228; PDZ; 1.
SQ   SEQUENCE   1822 AA;  201923 MW;  479821CE8820B4FD CRC64;

  Query Match             5.4%;  Score 112;  DB 11;  Length 1822;
  Best Local Similarity   20.0%;  Pred. No. 2.2;
  Matches   91;  Conservative   65;  Mismatches  152;  Indels  148;  Gaps   19;

Qy       2 NSTESNSSASVRPKVLSSIQEGTVPDNTSSARG--KEDVNKTLPNLQVVNHQQGPHHRHI 59
           ||| | : |:     :|| :    |   ||||   ||::: |  :
Db    1429 NSTFSINDATSHTSTMSS-RHSASPVVFSSARSSPKEELHPTTSS--------------- 1472

Qy      60 LKLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHF------------------------- 94
             :| ||  :     ||:   ||:  :|||     :
Db    1473 -QLAPSFSS-----SSSSSSGPRTFYPRQGATSKYLIGWKKPEGTINSVGFMDTRKRHQS 1526

Qy      95 DINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSM---TPT 151
           | ||              |   :| |:| |:  |:    : :|:  ||| :   :||
Db    1527 DGNE---------------IAHTRLRASTRDLRASPKPTSKSTIEE-DLKKLIDLESPT 1569

Qy     152 QNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCN----ASEREVAAFSNGYTADH- 206
             :|   :   ||:   :        :  : ::: | :    :|:||    |     |
Db    1570 PESQKNFKF---HGLSSPQSPFPSTPTSRRALHRTLSDESIYSSQREHFFTSRASLLDQA 1626

Qy     207 -------ERAYAAL-QHWTIRGPEASLA-----------QLISALRQHRRNDVVEKIRG 246
                 | :| :   :| |  :|                :: ::: ||  : :   |
Db    1627 LPNDVFFSSTYPSLPKSLPLRRPSYTLGMKSLHGEFFASDSSLTDIQETRRQPIPDP--G 1684

Qy     247 LMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESE- 305
           ||                    |:|   : |: | |:    : : :  ||    |
Db    1685 LM----------------------PLPDTASDLDWSNLVDAAKAYEVQRASFFAASDEN 1721

Qy     306 --PLLRCDSTSSGSSAL-----------SRNGSFITKEKKDTVLRQVRLDPCDLQPIF 350
             ||    | :| |  |             |:: |     | | :   :::   ||:
```

```
Db   1722 HRPL----SAASNSDQLEEQALVQMKSYSSSKDSSPTLASKVDQLEGMLKMLREDLKKEK 1777

Qy    351 DDMLHF------LNPEELRVIEEIPQAEDKLDRLFE 380
          :|  |       |  : ||: ||   | ||| :  |
Db   1778 EDKAHLQAEVEHLREDNLRLQEESQNASDKLKKFTE 1813


RESULT  14
O43166
ID   O43166      PRELIMINARY;      PRT;  1138 AA.
AC   O43166;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   KIAA0440 (FRAGMENT).
GN   KIAA0440.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Ishikawa K., Nagase T., Nakajima D., Seki N., Ohira M., Miyajima N.,
RA   Tanaka A., Kotani H., Nomura N., Ohara O.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB007900; BAA23712.1; -.
DR   InterPro; IPR000331; Rap_GAP.
DR   InterPro; IPR001478; PDZ.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF02145; Rap_GAP; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   SMART; SM00228; PDZ; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   1138 AA;  126014 MW;  CEE33A605029F325 CRC64;


  Query Match             5.3%;  Score 111;  DB 4;  Length 1138;
  Best Local Similarity   20.2%;  Pred. No. 1.5;
  Matches   92;  Conservative   59;  Mismatches  157;  Indels  148;  Gaps   19;

Qy      2 NSTESNSSASVRPKVLSSIQEGTVPDNTSSARG--KEDVNKTLPNLQVVNHQQGPHHRHI 59
          ||| | : |:     :|| :    |   :|||   ||:::   |:
Db    745 NSTFSINDAASHTSTMSS-RHSASPVVFTSARSSPKEELHPAAPS-------------- 788

Qy     60 LKLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHF------------------------ 94
           :| ||  :     ||:   ||:  :|||     :
Db    789 -QLAPSFSS-----SSSSSSGPRSFYPRQGATSKYLIGWKKPEGTINSVGFMDTRKRHQS 842

Qy     95 DINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSM---TPT 151
          | ||                |   :| |:| |:  |:    : :|:  ||| :   :||
Db    843 DGNE---------------IAHTRLRASTRDLRASPKPTSKSTIEE-DLKKLIDLESPT 885

Qy    152 QNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYA 211
           :|   :   | :   :        :  : ::: | : |          | :  |  :
Db    886 PESQKSFKF---HALSSPQSPFPSTPTSRRALHRTLSDESI---------YNSQREHFFT 933

Qy    212 ALQHWTIRGPEASLAQLISALRQHRRNDV----------------------VEKIRGLM 248
          :          |||       | |   |||                          :: : |
Db    934 S---------RASL------LDQALPNDVLFSSTYPSLPKSLPLRRPSYTLGMKSLHGEF 978

Qy    249 EDTTQLETDKLALPMSPSPLSPS--PIPSPNAKLENSALLTVEPSPQDKNKGFFVDESE- 305
          :    ||       | |  |   |:|   | |: | |:    : : :   ||    |
Db    979 SASDSSLTDIQETRRQPMP-DPGLMPLPDTAADLDWSNLVDAAKAYEVQRASFFAASDEN 1037

Qy    306 --PLLRCDSTSSGSSAL------------SRNGSFITKEKKDTVLRQVRLDPCDLQPIF 350
            ||    | :| |  |             |:: |     | | :   ::   ||:
Db   1038 HRPL----SAASNSDQLEDQALAQMKPYSSSKDSSPTLASKVDQLEGMLKMLREDLKKEK 1093

Qy    351 DDMLHF------LNPEELRVIEEIPQAEDKLDRLFE 380
          :|  |       |  : ||: ||   | ||| :  |
Db   1094 EDKAHLQAEVQHLREDNLRLQEESQNASDKLKKFTE 1129
```

```
RESULT  15
O95321
ID   O95321      PRELIMINARY;      PRT;  1783 AA.
AC   O95321;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE GAP PROTEIN ALPHA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99078010; PubMed=9858596;
RA   Gao Q., Srinivasan S., Boyer S.N., Wazer D.E., Band V.;
RT   "The E6 oncoproteins of high-risk papillomaviruses bind to a novel
RT   putative GAP protein, E6TP1, and target it for degradation.";
RL   Mol. Cell. Biol. 19:733-744(1999).
DR   EMBL; AF090989; AAD12543.1; -.
DR   InterPro; IPR000331; Rap_GAP.
DR   InterPro; IPR001478; PDZ.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF02145; Rap_GAP; 1.
DR   PROSITE; PS50106; PDZ; 1.
DR   SMART; SM00228; PDZ; 1.
SQ   SEQUENCE   1783 AA;  197493 MW;  8C9FB14DBD867FCC CRC64;


  Query Match             5.3%;  Score 111;  DB 4;  Length 1783;
  Best Local Similarity   20.2%;  Pred. No. 2.6;
  Matches   92;  Conservative   59;  Mismatches  157;  Indels  148;  Gaps   19;

Qy      2 NSTESNSSASVRPKVLSSIQEGTVPDNTSSARG--KEDVNKTLPNLQVVNHQQGPHHRHI 59
          ||| | : |:     :|| :    |   :|||   ||:::   |:
Db   1390 NSTFSINDAASHTSTMSS-RHSASPVVFTSARSSPKEELHPAAPS-------------- 1433

Qy     60 LKLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHF------------------------ 94
           :| ||  :     ||:   ||:  :|||     :
Db   1434 -QLAPSFSS-----SSSSSSGPRSFYPRQGATSKYLIGWKKPEGTINSVGFMDTRKRHQS 1487

Qy     95 DINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKAGLKKSM---TPT 151
          | ||                |   :| |:| |:  |:    : :|:  ||| :   :||
Db   1488 DGNE---------------IAHTRLRASTRDLRASPKPTSKSTIEE-DLKKLIDLESPT 1530

Qy    152 QNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSNGYTADHERAYA 211
           :|   :   | :   :        :  : ::: | : |          | :  |  :
Db   1531 PESQKSFKF---HALSSPQSPFPSTPTSRRALHRTLSDESI---------YNSQREHFFT 1578

Qy    212 ALQHWTIRGPEASLAQLISALRQHRRNDV----------------------VEKIRGLM 248
          :          |||       | |   |||                          :: : |
Db   1579 S---------RASL------LDQALPNDVLFSSTYPSLPKSLPLRRPSYTLGMKSLHGEF 1623

Qy    249 EDTTQLETDKLALPMSPSPLSPS--PIPSPNAKLENSALLTVEPSPQDKNKGFFVDESE- 305
          :    ||       | |  |   |:|   | |: | |:    : : :   ||    |
Db   1624 SASDSSLTDIQETRRQPMP-DPGLMPLPDTAADLDWSNLVDAAKAYEVQRASFFAASDEN 1682

Qy    306 --PLLRCDSTSSGSSAL------------SRNGSFITKEKKDTVLRQVRLDPCDLQPIF 350
            ||    | :| |  |             |:: |     | | :   ::   ||:
Db   1683 HRPL----SAASNSDQLEDQALAQMKPYSSSKDSSPTLASKVDQLEGMLKMLREDLKKEK 1738

Qy    351 DDMLHF------LNPEELRVIEEIPQAEDKLDRLFE 380
          :|  |       |  : ||: ||   | ||| :  |
Db   1739 EDKAHLQAEVQHLREDNLRLQEESQNASDKLKKFTE 1774


Search completed: January  4, 2002, 08:39:09
Job time: 137 sec
```

```
OM of: US-09-733-956-2 to: EST:*   out_format : pfs

Date: Jan 4, 2002  9:08 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09733956/runat_04012002_084024_15500/app_query.fasta_1.467
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09733956_@CGN1_1_3822 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-733-956-2
Query length: 405
Database: EST:*
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 1538.550000



score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation ..
gb_est2:BG819951       +  1218.50 1997.83  4.1e-102  848  ! BG819951 602782059F1 NCI_CGAP_B
gb_est2:BF681798       +  1177.50 1928.11  3.1e-98   1039  ! BF681798 602117463F1 Soares_ma
gb_est1:AU133604       +  1158.00 1899.81  1.2e-96   704  ! AU133604 AU133604 OVARC1 Homo s
gb_est2:BI102458       +  1108.00 1814.69  6.5e-92   911  ! BI102458 602889651F1 NCI_CGAP_K
gb_est1:AW211328       +   956.50 1567.01  4.0e-78   714  ! AW211328 uo79f09.y1 NCI_CGAP_Ma
gb_est2:BG656232       +   953.00 1563.70  6.2e-78   558  ! BG656232 ib38f11.y1 HR85 islet
gb_est2:BG243048       +   942.50 1542.02  1.0e-76   861  ! BG243048 602355338F1 NCI_CGAP_M
gb_est2:BG260064       +   925.00 1516.94  2.5e-75   589  ! BG260064 602371650F1 NIH_MGC_93
gb_est2:BF682316       -   913.00 1493.47  5.1e-74   848  ! BF682316 602117463T1 Soares_mam
gb_est1:AI552031       +   906.00 1486.29  1.3e-73   548  ! AI552031 mm72f10.y1 Stratagene
gb_est1:AV589747       +   866.00 1419.77  6.4e-70   575  ! AV589747 AV589747 Bos taurus br
gb_est2:BG923601       +   830.00 1357.52  1.9e-66   762  ! BG923601 602823467F1 NCI_CGAP_M
gb_est1:BE478092       +   697.00 1141.75  2.0e-54   521  ! BE478092 162175 BARC 5BOV Bos t
gb_est1:AA072902       +   668.00 1096.57  6.5e-52   398  ! AA072902 mm72f10.r1 Stratagene
gb_est2:BF129886       +   660.50 1078.65  6.4e-51   692  ! BF129886 601817687F1 NIH_MGC_58
gb_est1:AW385561       +   645.50 1056.35  1.1e-49   541  ! AW385561 RC3-LT0022-141299-011-
gb_est1:AA351536       -   635.00 1042.88  6.3e-49   368  ! AA351536 EST59310 Infant brain
gb_est1:AW963722       +   603.00  989.91  5.6e-46   373  ! AW963722 EST375795 MAGE reseque
gb_est1:BE485889       +   596.00  978.04  2.6e-45   385  ! BE485889 173172 BARC 5BOV Bos t
gb_est2:BE915474       +   589.00  960.13  2.6e-44   726  ! BE915474 601664973F1 NCI_CGAP_M
gb_est1:AA357231       +   588.00  964.96  1.4e-44   380  ! AA357231 EST66192 LNCAP cells I
gb_est2:BE936512       +   576.00  943.24  2.2e-43   460  ! BE936512 QV2-NT0048-160800-316-
gb_est1:BE184341       -   570.00  934.96  6.5e-43   391  ! BE184341 CM0-HT0676-010500-355-
gb_est2:BF090549       -   570.00  934.05  7.3e-43   428  ! BF090549 QV2-NT0048-040900-347-
gb_est2:BF090841       -   570.00  934.03  7.3e-43   429  ! BF090841 QV2-NT0048-140900-371-
gb_est2:BF090716       -   570.00  933.98  7.4e-43   431  ! BF090716 QV2-NT0048-120900-360-
gb_est2:BF090591       -   570.00  933.87  7.5e-43   436  ! BF090591 QV2-NT0048-050900-348-
gb_est1:BE763718       -   570.00  933.69  7.6e-43   444  ! BE763718 QV2-NT0048-150800-317-
gb_est2:BF090622       -   570.00  933.66  7.7e-43   445  ! BF090622 QV2-NT0048-050900-349-
gb_est2:BF090792       +   565.00  925.10  2.3e-42   459  ! BF090792 QV2-NT0048-140900-369-
gb_est2:BF090809       +   564.00  923.76  2.7e-42   445  ! BF090809 QV2-NT0048-140900-369-
gb_est1:AA155873       +   562.00  920.82  4.0e-42   429  ! AA155873 zo33a07.r1 Stratagene
gb_est2:BE936507       +   562.00  920.48  4.2e-42   444  ! BE936507 QV2-NT0048-160800-316-
gb_est2:BG287433       +   562.00  911.48  1.3e-41  1090  ! BG287433 602382024F1 NIH_MGC_9
gb_est1:BE763659       -   561.00  918.83  5.1e-42   444  ! BE763659 QV2-NT0048-150800-318-
gb_est1:BE763679       -   559.00  915.50  7.9e-42   445  ! BE763679 QV2-NT0048-150800-318-
gb_est2:BG777640       +   559.00  909.31  1.7e-41   826  ! BG777640 602665239F1 NIH_MGC_60
gb_est1:BE184340       -   557.00  913.92  9.7e-42   375  ! BE184340 CM0-HT0676-010500-355-
gb_est2:BF762074       -   553.00  908.24  2.0e-41   342  ! BF762074 IL2-CS0048-181000-196-
gb_est2:BF090585       +   550.00  900.58  5.3e-41   448  ! BF090585 QV2-NT0048-050900-348-
gb_est2:BF090621       +   542.00  887.28  2.9e-40   452  ! BF090621 QV2-NT0048-050900-349-
gb_est1:BE184321       -   541.00  886.15  3.4e-40   429  ! BE184321 CM0-HT0676-010500-3
gb_est1:AW784943       +   536.00  876.31  1.2e-39   503  ! AW784943 115287 MARC 1PIG Su
gb_est2:D59902         +   535.00  878.20  9.4e-40   353  ! D59902 HUM073E08B Clontech h
gb_est2:BF090710       +   532.00  871.02  2.4e-39   441  ! BF090710 QV2-NT0048-120900-3

seq_name: gb_est2:BG819951

seq_documentation_block:
LOCUS       BG819951      848 bp    mRNA            EST       22-MAY-2001
DEFINITION  602782059F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4932855
            5', mRNA sequence.
ACCESSION   BG819951
VERSION     BG819951.1  GI:14167538
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 848)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10859  row: e  column: 16
            High quality sequence stop: 745.
FEATURES             Location/Qualifiers
     source          1. .848
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4932855"
                     /clone_lib="NCI_CGAP_Brn67"
                     /tissue_type="anaplastic oligodendroglioma with 1p/19q
                     loss"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.3 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      221 a    236 c    243 g    148 t
ORIGIN


alignment_scores:
            Quality: 1218.50              Length:    261
              Ratio:  4.816                Gaps:      1
 Percent Similarity:  96.935   Percent Identity: 93.870

alignment_block:
 US-09-733-956-2 x BG819951   ..

 Align seg 1/1  to: BG819951  from: 1  to: 848

      25 ValProAspAsnThrSerSerAlaArgGlyLysGluAspValAsnLysTh 41
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       2 GTCCCTGACAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGAC 51

      41 rLeuProAsnLeuGlnValValAsnHisGlnGlnGlyProHisHisArgH 58
         |||||||:::||||||||||||||||||||||||||||||||||||||||
      52 CCTCCCAAAGCTTCAGGTAGTCAACCACCAGCAAGGCCCCCACCACAGAC 101

      58 isIleLeuLysLeuLeuProSerMetGluAlaThrGlyGlyGluLysSer 74
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     102 ACATCCTGAAGCTGCTGCCGTCCATGGAGGCCACTGGGGGCGAGAAGTCC 151

      75 SerThrProIleLysGlyProLysArgGlyHisProArgGlnAsnLeuHi 91
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 152 AGCACGCCCATCAAGGGCCCCAACAGGGGACATCCTAGACAGAACCTACA 201

  91 sLysHisPheAspIleAsnGluHisLeuProTrpMetIleValLeuPheL 108
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 202 CAAGCATTTTGACATCAATGAGCATTTGCCCTGGATGATTGTGCTTTTCC 251

 108 euLeuLeuValLeuValValIleValValCysSerIleArgLysSerSer 124
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 252 TGCTGCTGGTGCTTGTGGTGATTGTGGTGTGCAGTATCCGGAAAAGCTCG 301

 125 ArgThrLeuLysLysGlyProArgGlnAspProSerAlaIleValGluLy 141
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 302 AGGACTCTGAAAAAGGGGCCCCGGCAGGATCCCAGTGCCATTGTGGAAAA 351

 141 sAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgGluLysTrpI 158
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 352 GGCAGGGCTGAAGAAATCCATGACTCCAACCCAGAACCGGGAGAAATGGA 401

 158 leTyrTyrCysAsnGlyHisGlyIleAspIleLeuLysLeuValAlaAla 174
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 402 TCTACTACTGCAATGGCCATGGTATCGATATCCTGAAGCTTGTAGCAGCC 451

 175 GlnValGlySerGlnTrpLysAspIleTyrGlnPheLeuCysAsnAlaSe 191
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 452 CAAGTGGGAAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAG 501

 191 rGluArgGluValAlaAlaPheSerAsnGlyTyrThrAlaAspHisGluA 208
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 502 TGAGAGGGAGGTTGCTGCTTTCTCCAATGGGTACACAGCCGACCACGAGC 551

 208 rgAlaTyrAlaAlaLeuGlnHisTrpThrIleArgGlyProGluAlaSer 224
     ||||||||||||||||||||||||||||||||   ||||||||||||||||
 552 GGGCCTACGCAGCTCTGCAGCACTGGACCATCCG.GGGCCCGAGGCCAGC 600

 225 LeuAlaGlnLeuIleSerAlaLeuArgGlnHisArgArgAsnAspValVa 241
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 601 CTCGCCCAGCTAATTAGCGCCCTGCGCCAGCACCGGAGAAACGATGTCGT 650

 241 lGluLysIleArgGlyLeuMetGluAspThrThrGlnLeuGluThrAspL 258
     ||||||||||   ||||||||||||||||||:::||||||||||||||||
 651 GGAGAAGATTCCGGGGCTGATGGAAGACACCAGCCAGCTGGAAACTGACA 700

 258 ysLeuAlaLeuProMetSerProSerProLeuSerProSerProIlePro 274
     ||:::|||   ||||||||||||   |||||||:::   :::|||   |||
 701 AAATAGCTCT.CCGATGAGCCCCAGGCCGTTAGCC...GAGCCCAATCCC 746

 275 SerProAsnAlaLysLeuGluAsnSerAlaLeu 285
     ||||||:::     ||||:::   :::|||
 747 AGCCCGACGCGAACTTGAGAATCGCGATCTCTG 779

seq_name: gb_est2:BF681798

seq_documentation_block:
LOCUS       BF681798     1039 bp    mRNA            EST       21-DEC-2000
DEFINITION  602117463F1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:3468855 5', mRNA sequence.
ACCESSION   BF681798
VERSION     BF681798.1  GI:11955693
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1039)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: M. Bento Soares, Ph.D.
             cDNA Library Preparation: Bento Soares and M. Fatima Bonaldo
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8476  row: m  column: 16
            High quality sequence start: 6
            High quality sequence stop: 771.
FEATURES             Location/Qualifiers
     source          1. .1039
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3468855"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      274 a    306 c    280 g    179 t
ORIGIN


alignment_scores:
            Quality: 1177.50              Length:    345
              Ratio:  4.282                 Gaps:      8
 Percent Similarity:  79.710   Percent Identity: 72.174

alignment_block:
 US-09-733-956-2 x BF681798   ..

 Align seg 1/1  to: BF681798  from: 1  to: 1039

  35 LysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHisGl 51
     ||||||   :::|||:::|||||||||||          |||:::|||||
   4 AAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCACCA 53

  51 nGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMetG 67
     ||||:::|||||||||||||||||||||||||||||||   |||||||
  54 GCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCATGG 103

  67 luAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArg 83
     ||||||||  ||||||||||||||||||   ||||||:::||||||||||
 104 AGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGAGG 150

  84 GlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLe 100
     ||||||||||||||||||   ||||||||||||||||||||||||||||
 151 GGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCACTT 200

 100 uProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValV 117
     ||||||||||||||||||||||||||||||||||||||:::|||||||
 201 GCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGTGG 250

 117 alCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 133
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 251 TGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGCAG 300

 134 AspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPr 150
     ||||||||||||||||||||||||||||||||||||||||:::||||||
 301 GATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACTCC 350

 150 oThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleA 167
     ||||||||||||||||||||||||||||||   ||||||||||||||||
 351 CACTCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTATTG 400

 167 spIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIle 183
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 401 ACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACATC 450
```

```
     184 TyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAs 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 TATCAGTTTCTTTGCAACGCCAGCGAGAGGGAGGTGGCGGCCTTCTCCAA 500

     200 nGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpT 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 TGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCAGCACTGGA 550

     217 hrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArg 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     551 CCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGCGC 600

     234 GlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAs 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGA 650

     250 pThrThrGlnLeuGluThrAspLysLeuAlaLeu.ProMetSerProSer 266
         |||||||||||||||||||||||||||||||||| ||||||||||||||||
     651 CACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCAATGAGCCCCAGT 700

     267 ProLeuSerProSerPro.IleProSerProAsnAlaLysLeuGluAsnS 283
         ||||||||||||||||||       ||||||||||:::|||||||||||||
     701 CCGCTTAGCCCGAGCCCATGCC...AGTCCTAACGTGAAACTTGAGAATT 747

     283 erAlaLeuLeuThrValGluPro..............................  290
         ||   ||||||||||||||||||
     748 CCCATCTCCTGACAGTGGAGCCTCACGCTGGGACAGAACAAGTGCTCTCG 797

     291 .....................................SerProGlnAspLy 295
                                              :::   |||
     798 GGGACGAGTCAAACCCTTTGGGAAGCGATCCCATCAGGGTCTCAGACTGG 847

     295 sAsnLysGlyPhePheValAspGluSerGlu..................... 305
          |||      ::::::      :::::::::
     848 AAACGGTCCTTATTACAAAAAACAGGGACAGGTGCGAGAGCGCGGACCGG 897

     306 ........ProLeuLeuArgCysAspSerThrSerSerGlySerSerAla 319
                 |||:::|||      :::   |||   ||||||::::::
     898 TGAATGGCCCCATGTTGACGGTGAACCTGACCCGAAGCGGTCAGAACCGG 947

     320 ..................LeuSerArgAsnGlySerPheIleThrLysGl 330
                           :::::::|||:::           |||   ::
     948 GGAAAAGCCCCAAAAGATGTCACACGAACCACCTTGTAGTACACATAAGA 997

     330 uLysLysAspThrValLeuArgGlnValArg 340
         :   ::::::|||:::   |||   :::|||
     998 CGACGAACAGACGATCCCACGCGCTATACGG 1028
```

seq_name: gb_est1:AU133604

seq_documentation_block:

```
LOCUS       AU133604      704 bp    mRNA            EST       24-OCT-2000
DEFINITION  AU133604 OVARC1 Homo sapiens cDNA clone OVARC1000267 5', mRNA
            sequence.
ACCESSION   AU133604
VERSION     AU133604.1  GI:10994143
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 704)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .704
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="OVARC1000267"
                     /clone_lib="OVARC1"
                     /tissue_type="ovary, tumor tissue"
                     /note="Vector: pME18SFL3"
BASE COUNT      176 a    204 c    189 g    132 t      3 others
ORIGIN
```

alignment_scores:

```
            Quality: 1158.00            Length:    234
              Ratio:  5.035               Gaps:      0
 Percent Similarity:  98.291   Percent Identity: 97.863
```

alignment_block:

US-09-733-956-2 x AU133604 ..

Align seg 1/1 to: AU133604 from: 1 to: 704

```
     115 IleValValCysSerIleArgLysSerSerArgThrLeuLysLysGlyPr 131
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       3 ATTGTGGTGTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCC 52

     131 oArgGlnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerM 148
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      53 TCGGCAGGATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCA 102

     148 etThrProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHis 164
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 TGACTCCAACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCAT 152

     165 GlyIleAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLy 181
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     153 GGTATCGATATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAA 202

     181 sAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaP 198
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     203 AGATATCTATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTT 252

     198 heSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGln 214
         ||||||||||||||||||||||||:::||||||||||||||||||||||
     253 TCTCCAATGGGTACACAGCCGACAACGAGCGGGCCTACGCAGCTCTGCAG 302

     215 HisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAl 231
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     303 CACTGGACCATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGC 352

     231 aLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuM 248
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     353 CCTGCGCCAGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGA 402

     248 etGluAspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSer 264
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     403 TGGAAGACACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGC 452

     265 ProSerProLeuSerProSerProIleProSerProAsnAlaLysLeuGl 281
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     453 CCCAGCCCGCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGA 502

     281 uAsnSerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysG 298
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     503 GAATTCCGCTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGG 552

     298 lyPhePheValAspGluSerGluProLeuLeuArgCysAspSerThrSer 314
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      553 GCTTCTTCGTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCC 602

      315 SerGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLy 331
          ||||||||||||||||||||   ||||||||||||||||||||||||||
      603 AGCGGCTCCTCCGCGCTGAGCANGAACGGTTCCTTTATTACCAAAGAAAA 652

      331 sLysAspThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnP 348
          |||||||||||||||   ||||||   ||||||||||||||   ||||
      653 GAAGGACACAGTGTTGCNGCAAGTACC.CTGGACCCCTGTGACTTNCAAC 701

      348 ro 348
          ||
      702 CT 703

seq_name: gb_est2:BI102458

seq_documentation_block:
LOCUS       BI102458      911 bp    mRNA            EST       26-JUN-2001
DEFINITION  602889651F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044620
            5', mRNA sequence.
ACCESSION   BI102458
VERSION     BI102458.1  GI:14553351
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11122  row: f  column: 13
            High quality sequence stop: 736.
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5044620"
                     /clone_lib="NCI_CGAP_Kid14"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT      211 a    291 c    240 g    169 t
ORIGIN


alignment_scores:
            Quality: 1108.00            Length:    282
              Ratio:  4.362              Gaps:      5
 Percent Similarity:  90.071   Percent Identity: 84.043

alignment_block:
 US-09-733-956-2 x BI102458   ..

 Align seg 1/1  to: BI102458  from: 1  to: 911

      103 MetIleValLeuPheLeuLeuLeuValLeuValValIleValValCysSe 119
          ||||||||||||||||||||||||||||||||:::|||||||||||||||
        2 ATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGTGGTGTGCAG 51

      119 rIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnAspProS 136
          ||||||||||||||||||||||||||||||||||||||||||||||||||
       52 TATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGCAGGATCCCA 101

      136 erAlaIleValGluLysAlaGlyLeuLysLysSerMetThrProThrGln 152
          ||||||||||||||||||||||||||||||||:::||||||||||||||
      102 GCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACTCCCACCCAG 151

      153 AsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAspIleLe 169
          ||||||||||||||||||||||   ||||||||||||||||||||||||
      152 AACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTATTGACATCTT 201

      169 uLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleTyrGlnP 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      202 GAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACATCTATCAGT 251

      186 heLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyr 202
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      252 TTCTTTGCAACGCCAGCGAGAGGGAGGTGGCGGCCTTCTCCAATGGATAC 301

      203 ThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpThrIleAr 219
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      302 ACTGCAGATCATGAACGGGCCTACGCGGCTCTGCAGCACTGGACCATCCG 351

      219 gGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgGlnHisA 236
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      352 TGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGCGCCAGCACC 401

      236 rgArgAsnAspValValGluLysIleArgGlyLeuMetGluAspThrThr 252
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      402 GACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGACACCACG 451

      253 GlnLeuGluThr.AspLysLeuAlaLeuProMetSerProSerProLeuS 269
          ||||||||||| |||||||||||||||||||||||||||||||||||||||
      452 CAGTTGGAAACCAGACAAACTGGCTCTCCCCATGAGCCCCAGTCCGCTGA 501

      269 erProSerProIleProSerProAsnAlaLysLeuGluAsnSerAlaLeu 285
          |||||||||||||||||||||||:::|||||||||||||||:::|||
      502 GCCCGAGCCCCATCCCCAGTCCTAACGTGAAACTTGAGAATTCCACTCTC 551

      286 LeuThrValGluProSerProGlnAsp.LysAsnLysGlyPhePheValA 302
          ||||||||||||||||||||   ||| |||||||||   |||||||||
      552 CTGACAGTGGAGCCCTCACCGCTGGACAAAGAACAAGTGCTTCTTCGTGG 601

      302 spGluSerGluProLeuLeuArgCysAspSerThr.SerSerGlySerSe 318
          |||||||||||||||||||||||||||||||||||| ||||||||||||||
      602 ACGAGTCAGAGCCCCTTCTGCGTTGCGACTCCACCATCCAGTGGCTCTTC 651

      318 rAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspThrV 335
          |||||||   ::::::   |||||||||||:::|||:::|||
      652 AGCACTGAGGCAGAAACGGTCCTTTATTACCAACGAACAGAAGACACAGT 701

      335 alLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePheAsp 351
                :::             ||||||||||||||||||   ||||||
      702 GTTGCGGCAGG...TCGGCTGGACCCTGTGACTTGCAGCCATC.TTTGAT 747

      352 AspMetLeuHisPheLeu...AsnProGluGluLeuArgValIleGluGl 367
          |||||||||   :::       :::|||||||||||||||      :::::
      748 GACATGCTGCCATATCCCTGAACCCCCGAGGAGCTGCGGTGATCGAACGA 797

      367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGlu 380
          :      :::||||||::::::|||      |||   :::
      798 CGACTCCCATGCTGAGGAACAACTGGCCGGCCTTCACGAT 837

seq_name: gb_est1:AW211328

seq_documentation_block:
LOCUS       AW211328      714 bp    mRNA            EST       03-DEC-1999
DEFINITION  uo79f09.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:2648777 5'
            similar to TR:O75509 O75509 TNFR-RELATED DEATH RECEPTOR-6. ;, mRNA
            sequence.
ACCESSION   AW211328
VERSION     AW211328.1  GI:6517276
```

```
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 714)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: uo79f09.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            MGI:1029229
            Seq primer: -40RP from Gibco
            High quality sequence stop: 426.
FEATURES              Location/Qualifiers
     source           1. .714
                      /organism="Mus musculus"
                      /strain="129 - C57/B6 - FVBN"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:2648777"
                      /clone_lib="NCI_CGAP_Mam3"
                      /tissue_type="tumor, gross tissue"
                      /dev_stage="10 months"
                      /lab_host="DH10B"
                      /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Library constructed by Life Technologies. Investigators
                      providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                      Reference for transgenic model: Xu et al., Nature Genetics
                      22, 37-43 (1999)."
BASE COUNT      186 a    215 c    184 g    128 t      1 others
ORIGIN


alignment_scores:
            Quality:  956.50              Length:    237
              Ratio:  4.470                 Gaps:      6
 Percent Similarity:  90.295   Percent Identity: 84.388

alignment_block:
 US-09-733-956-2 x AW211328   ..

 Align seg 1/1  to: AW211328  from: 1  to: 714

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||:::|||||||||:::|||||||||||   ||||||
      13 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 62

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
          |||:::|||:::|||||||||||||||||||||||||||||:::   |
      63 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 112

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||   :::|||:::||||||||||||         |||:::|||
     113 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 162

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
         ||||||:::|||||||||||||||||||||||||||||||   |||||
     163 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 212

      66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
         ||||||||||   ||||||||||||||||||   ||||||:::|||||||
     213 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 259

      83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
         ||||||||||||||||||||   ||||||||||||||||||||||||||||
     260 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 309

     100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
         ||||||   ||||||||||||||||||||||||||||||||||:::|||||
     310 TTGCCCTTGATGATCGTTCTCTTTCTTCTGCTTGTCCTGGTGCTGATAGT 359

     116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     360 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 409

     133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
         ||||||||||||||||||||||||||||||||||||||||||:::|||
     410 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 459

     150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
         ||||||||||||||||||||||||||||||   ||||||||||||||
     460 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 509

     166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     510 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACA 559

     183 leTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
         ||||||||||||||||||||||||||||   ||||||||||||||||||
     560 TCTATCAGTTTCTTTGCAACGCCAGCGAAGACGAGGTGGCGGCCTTCTCC 609

     200 AsnGlyTyrThrAlaAspHisGlu.ArgAlaTyrAlaAlaLeu.GlnHis 215
         ||||||||||||   :::||||||| ||||||||||||||||| ||||||
     610 AATGGATACACTTGAAATCACGAACCGGGCCTACGCNGCTCTGGCAGCAC 659

     216 .TrpThrIleArgGlyPro...GluAlaSerLeuAlaGlnLeuIleSerA 231
          ||||||||||||||||||         :::|||   ||||||||||||
     660 CTGGACCATTCGTGGGCCCTGAGGCCCAGCCTTGCCCCAGCTTATTAGCC 709

     231 laLeu 232
           |||
     710 GCCTG 714

seq_name: gb_est2:BG656232

seq_documentation_block:
LOCUS       BG656232      558 bp    mRNA            EST       05-JUL-2001
DEFINITION  ib38f11.y1 HR85 islet Homo sapiens cDNA 5' similar to TR:O75509
            O75509 TNFR-RELATED DEATH RECEPTOR-6. ;, mRNA sequence.
ACCESSION   BG656232
VERSION     BG656232.1  GI:13793641
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
```

```
          obtaining a clone please contact: Dr. Hiroshi Inoue
          (hinoue@im.wustl.edu)
          Seq primer: -40RP from Gibco
          High quality sequence stop: 475.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HR85 islet"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
BASE COUNT      127 a    172 c    148 g    111 t
ORIGIN


alignment_scores:
            Quality:  953.00            Length:    186
              Ratio:  5.124               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.462

alignment_block:
 US-09-733-956-2 x BG656232   ..

 Align seg 1/1  to: BG656232  from: 1  to: 558

     181 LysAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAl 197
         :::||||||||||||||||||||||||||||||||||||||||||||||||
       1 GAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGC 50

     197 aPheSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuG 214
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 TTTCTCCAATGGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGC 100

     214 lnHisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSer 230
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 AGCACTGGACCATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGC 150

     231 AlaLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLe 247
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GCCCTGCGCCAGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCT 200

     247 uMetGluAspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetS 264
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 GATGGAAGACACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGA 250

     264 erProSerProLeuSerProSerProIleProSerProAsnAlaLysLeu 280
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 GCCCCAGCCCGCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTT 300

     281 GluAsnSerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLy 297
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GAGAATTCCGCTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAA 350

     297 sGlyPhePheValAspGluSerGluProLeuLeuArgCysAspSerThrS 314
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 GGGCTTCTTCGTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACAT 400

     314 erSerGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGlu 330
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 CCAGCGGCTCCTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAA 450

     331 LysLysAspThrValLeuArgGlnValArgLeuAspProCysAspLeuGl 347
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 AAGAAGGACACAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCA 500

     347 nProIlePheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgV 364
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     501 GCCTATCTTTGATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGG 550

     364 alIleGlu 366
         ||||||||
     551 TGATTGAA 558

seq_name: gb_est2:BG243048

seq_documentation_block:
LOCUS       BG243048      861 bp    mRNA            EST       13-FEB-2001
DEFINITION  602355338F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4483713 5',
            mRNA sequence.
ACCESSION   BG243048
VERSION     BG243048.1  GI:12752863
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 861)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10322  row: k  column: 10
            High quality sequence stop: 712.
FEATURES             Location/Qualifiers
     source          1. .861
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4483713"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="3 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      226 a    259 c    221 g    155 t
ORIGIN


alignment_scores:
            Quality:  942.50            Length:    259
              Ratio:  4.116               Gaps:      10
 Percent Similarity:  88.417   Percent Identity: 80.695

alignment_block:
 US-09-733-956-2 x BG243048   ..

 Align seg 1/1  to: BG243048  from: 1  to: 861

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||:::||||||||||:::||||||||||||   ||||||
      79 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 128

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
          |||:::|||:::|||||||||||||||||||||||||||||:::   |
     129 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 178
```

```
 34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
    ||||||||   :::|||:::||||||||||||         |||:::|||
179 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 228

 51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
    ||||||:::|||||||||||||||||||||||||||||||   |||||
229 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 278

 66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
    |||||||||   ||||||||||||||||||   |||||||:::|||||||
279 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 325

 83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
    ||||||||||||||||||   |||||||||||||||||||||||||||||
326 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 375

100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
    ||||||||||||||||||||||||||||||||||||||||||:::|||||
376 TTGCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGT 425

116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
    ||||||||||||||||||||||||||||||||||||||||||||||||||
426 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 475

133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
    ||||||||||||||||||||||||||||||||||||||||||||:::|||
476 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 525

150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
    ||||||||||||||||||||||||||||||   ||||||||||||||||
526 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 575

166 eAspIleLeuLysLeuValAlaAlaGlnValGly..SerGlnTrpLysAs 182
    ||||||||||||||||||||||||||||||||  :::||||||||||||
576 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAAGCCCAGTGGAAGGA 625

182 pIleTyrGlnPheLeuCysAsnAlaSerGluArg.GluValAlaAlaPhe 198
    |||||||||||||||||||||||||||||||| ||||||||||||||||
626 CATCTATCAGTTTCTTTGCAACGCCAGCGAGAGGGGAGGTGGCGGCTTTC 675

199 Ser.AsnGlyTyrThr.AlaAspHisGluArgAlaTyrAla..AlaLeuG 214
    ||| ||||||||||||| ||||||||||||||||||||||||  |||||||
676 TCCAAATGGATACACTTGCAGATCATGAACGGGCCTACGCCGGGCTCTGC 725

214 lnHisTrp.ThrIle.ArgGlyProGluAlaSerLeuAlaGlnLeuIleS 230
    |||||||| |||||| |||||||||   :::   ||||||||||||
726 AGCACTGGAACCATCCCGTGGGCCCTGAGGCCACCTTGCCCAGCTCAATT 775

230 erAlaLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGly 246
      :::|||         ||||||:::::::::   ||||||:::   |||
776 AAGGCCTTGGGCCACCCCGACGCAAGAATTTGTGTGAAAAACTC...GGG 822

247 LeuMetGluAspThr 251
      :::|||   |||
823 GCGTTGGAAAAAACC 837
```

seq_name: gb_est2:BG260064

seq_documentation_block:

```
LOCUS       BG260064      589 bp    mRNA            EST       13-FEB-2001
DEFINITION  602371650F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4479684 5',
            mRNA sequence.
ACCESSION   BG260064
VERSION     BG260064.1  GI:12769880
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10312  row: c  column: 13
            High quality sequence start: 7
            High quality sequence stop: 589.
FEATURES             Location/Qualifiers
     source          1. .589
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4479684"
                     /clone_lib="NIH_MGC_93"
                     /tissue_type="transitional cell papilloma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      146 a    155 c    172 g    116 t
ORIGIN
```

alignment_scores:

```
           Quality:  925.00            Length:    195
             Ratio:  4.894               Gaps:      4
 Percent Similarity:  96.923   Percent Identity: 95.385
```

alignment_block:

US-09-733-956-2 x BG260064 ..

Align seg 1/1 to: BG260064 from: 1 to: 589

```
 53 GlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGluAlaTh 69
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  3 GGTCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGAGGCCAC 52

 69 rGlyGlyGluLysSerSerThrProIleLysGlyProLys.ArgGlyHis 85
    |||||||||||||||||||||||||||||||||||||| |||||||||
 53 TGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGTAGGGGACAT 102

 86 ProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeuProTr 102
    ||||||||||||||||||||||||||||||||||||||||||||||||||
103 CCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTGCCCTG 152

102 p.MetIleValLeuPheLeuLeuLeuValLeuValValIleValValCys 118
    | ||||||||||||||||||||||||||||||||||||||||||||||||
153 GTATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGTGTGC 202

119 SerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnAspPr 135
    ||||||||||||||||||||||||||||||||||||||||||||||||||
203 AGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGGATCC 252

135 oSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrProThrG 152
    ||||||||||||||||||||||||||||||||||||||||||||||||||
253 CAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCAACCC 302

152 lnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAspIle 168
    ||||||||||||||||||||||||||||||||||||||||||||||||||
303 AGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGATATC 352

169 LeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleTyrGl 185
    ||||||||||||||||||||||||||||||||||||||||||||||||||
353 CTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCTATCA 402

185 nPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsnGlyT 202
    ||||||||||:::||||||||||||||||||||||||||||||||||||
```

```
     403 GTTTCTTTGCGATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAATGGGT 452

     202 yrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpThrIle 218
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     453 ACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGACCATC 502

     219 ArgGlyProGluAlaSerLeu..AlaGlnLeuIleSerAlaLeuArgGln 234
         ||||||||||||||||||||     ::::::||||||||||||||||||
     503 CGGGGCCCCGAGGCCAGCCTCGACCCAGCGTAATTAGCGCCCTGCGCCAG 552

     235 .HisArgArgAsnAspValValGluLys 243
          |||||||||||||||||||||||||
     553 TGACCGGAGAAACGATGTTGTGGAGAAA 580

seq_name: gb_est2:BF682316

seq_documentation_block:
LOCUS       BF682316      848 bp    mRNA            EST       21-DEC-2000
DEFINITION  602117463T1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:3468855 3', mRNA sequence.
ACCESSION   BF682316
VERSION     BF682316.1  GI:11956211
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 848)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: M. Bento Soares, Ph.D.
             cDNA Library Preparation: Bento Soares and M. Fatima Bonaldo
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8476  row: m  column: 16
            High quality sequence start: 21
            High quality sequence stop: 715.
FEATURES             Location/Qualifiers
     source          1. .848
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3468855"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      242 a    200 c    232 g    174 t
ORIGIN


alignment_scores:
            Quality:  913.00            Length:    207
              Ratio:  4.682               Gaps:      1
 Percent Similarity:  94.203  Percent Identity: 88.889

alignment_block:
 US-09-733-956-2 x BF682316/rev  ..

 Align seg 1/1  to reverse of: BF682316  from: 1  to: 848
```

```
     200 AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
         :::|||      :::   |||:::   |||   |||:::||||||   ||
     843 CATGGGTTACTCTCCGTTCATGACGGCGCTAACGCTGTTCTGCAGCCTTG 794

     216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeu. 232
         |||||||||||||||||||   :::||||||||||||||||||:::|||
     793 GACCATCCGTGGCCCTGAGCCAACCCTTGCCCAGCTCATTAGCTCCTTGC 744

     233 ArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGl 249
         :::||||||   ||||||||||||||||||||||||||||||||||||||
     743 CCAAGGCACCGAGGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGA 694

     249 uAspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProS 266
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     693 AGACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAGCCCCA 644

     266 erProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsn 282
         ||||||||||||||||||:::||||||||||:::||||||||||||||||
     643 GTCCGCTTAGCCCGAGCCCCATGCCCAGTCCTAACGTGAAACTTGAGAAT 594

     283 SerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPh 299
         |||:::|||||||||||||||||||||   ||||||||||   ||
     593 TCCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAGTGCTT 544

     299 ePheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerG 316
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     543 CTTCGTGGACGAGTCAGAGCCCCTTCTGCGATGCGACTCCACATCCAGTG 494

     316 lySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLys 332
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     493 GCTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAAGAAAAGAAG 444

     333 AspThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIl 349
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     443 GACACAGTGTTGCGGCAGGTCCGCCTGGACCCCTGTGACTTGCAGCCCAT 394

     349 ePheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleG 366
         ||||||||||||||||||:::||||||||||||||||||||||||||||||
     393 CTTTGATGACATGCTGCATATCCTGAACCCCGAGGAGCTGCGGGTGATTG 344

     366 luGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIle 382
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     343 AAGAGATTCCCCAGGCTGAGGACAAACTGGACCGCCTCTTCGAGATCATT 294

     383 GlyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSe 399
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     293 GGGGTCAAGAGCCAAGAAGCCAGCCAGACCCTCTTGGACTCTGTGTACAG 244

     399 rHisLeuProAspLeuLeu 405
         ||||||||||||||||||
     243 TCATCTTCCTGACCTATTG 225

seq_name: gb_est1:AI552031

seq_documentation_block:
LOCUS       AI552031      548 bp    mRNA            EST       15-MAR-2000
DEFINITION  mm72f10.y1 Stratagene mouse macrophage (#937306) Mus musculus cDNA
            clone IMAGE:533995 5', mRNA sequence.
ACCESSION   AI552031
VERSION     AI552031.1  GI:4484394
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 548)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This read is a RESEQUENCE of a previously sequenced mouse clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            MGI:320931
            Seq primer: -40RP from Gibco
            High quality sequence stop: 402
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:533995"
                     /clone_lib="Stratagene mouse macrophage (#937306)"
                     /tissue_type="macrophage"
                     /dev_stage="WEHI-3 cell line"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: blood; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      131 a    167 c    143 g    107 t
ORIGIN


alignment_scores:
           Quality:  906.00            Length:    182
             Ratio:  5.061               Gaps:      0
 Percent Similarity:  98.352   Percent Identity: 96.154

alignment_block:
 US-09-733-956-2 x AI552031    ..

 Align seg 1/1  to: AI552031  from: 1  to: 548

     148 MetThrProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHi 164
         :::||||||||||||||||||||||||||||||||||||   ||||||||
       2 CTGACTCCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCA 51

     164 sGlyIleAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpL 181
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      52 TGGTATTGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGA 101

     181 ysAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAla 197
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     102 AGGACATCTATCAGTTTCTTTGCAACGCCAGCGAGAGGGAGGTGGCGGCC 151

     198 PheSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGl 214
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     152 TTCTCCAATGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCA 201

     214 nHisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerA 231
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     202 GCACTGGACCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCG 251

     231 laLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeu 247
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     252 CCTTGCGCCAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTG 301

     248 MetGluAspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSe 264
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     302 ATGGAAGACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAG 351

     264 rProSerProLeuSerProSerProIleProSerProAsnAlaLysLeuG 281
         ||||||||||||||||||||||||:::|||||||||||:::|||||||
     352 CCCCAGTCCGCTTAGCCCGAGCCCCATGCCCAGTCCTAACGTGAAACTTG 401

     281 luAsnSerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLys 297
         |||||||:::||||||||||||||||||||||   ||||||||||||
     402 AGAATTCCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAG 451

     298 GlyPhePheValAspGluSerGluProLeuLeuArgCysAspSerThrSe 314
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     452 TGCTTCTTCGTGGACGAGTCAGAGCCCCTTCTGCGATGCGACTCCACATC 501

     314 rSerGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLys 329
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     502 CAGTGGCTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAA 547

seq_name: gb_est1:AV589747

seq_documentation_block:
LOCUS       AV589747      575 bp    mRNA            EST       29-AUG-2000
DEFINITION  AV589747 Bos taurus brain fetus Bos taurus cDNA clone E1BR008C08
            5', mRNA sequence.
ACCESSION   AV589747
VERSION     AV589747.1  GI:9700740
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 575)
  AUTHORS   Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
            Suzuki,H.
  TITLE     bovine cDNA sequencing
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES             Location/Qualifiers
     source          1. .575
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="E1BR008C08"
                     /clone_lib="Bos taurus brain fetus"
                     /tissue_type="brain"
                     /dev_stage="fetus"
                     /lab_host="DH10B"
                     /note="Vector: pZL1; Site_1: Sal1; Site_2: Not1; Poly A
                     was deleted from a Not1 site"
BASE COUNT      151 a    166 c    159 g     97 t      2 others
ORIGIN


alignment_scores:
           Quality:  866.00            Length:    189
             Ratio:  4.838               Gaps:      2
 Percent Similarity:  94.709   Percent Identity: 91.005

alignment_block:
 US-09-733-956-2 x AV589747    ..

 Align seg 1/1  to: AV589747  from: 1  to: 575

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         :::||||||||||||||||||||||||||||||||||||||||||||||||
      13 ATCAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTGCC 62

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
```

```
      |||:::   ||||||||||||||||||||||||||||||||||||
   63 AAGTGGCACCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCGAGCA 112

   34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
       :::|||  ||||||||||:::|||:::||||||||:::||||||
  113 GGGAGGAGGGCGTGAACAAGACCTTCCCCAAGCTCCAGGTAGCGAACCAC 162

   51 GlnGlnGlyProHisHis.ArgHisIleLeuLysLeuLeuProSerMetG 67
      |||||||||||||||||   ||||||||||||||||||||||||||||
  163 CAGCAAGGGCCCCACCACNNTACACATCCTGAAGCTGCTGCCGTCCATGG 212

   67 luAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArg 83
      |||||||||||||||||||||||||||||||||||||||    ||||||
  213 AGGCCACTGGGGGTGAGAAGTCAAGCACGCCCATCAAAGGCCTCAAGAGG 262

   84 GlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLe 100
      |||||||||||||||||    ||||||||||||||||||||||||||||
  263 GGCCACCCCAGGCAGAACCCACACAAGCATTTTGACATCAACGAGCACCT 312

  100 uProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValV 117
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  313 TCCGTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTGGTGGTGATAGTGG 362

  117 alCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGln 133
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  363 TGTGCAGTATCCGGAAAAGCTCGAGGACTCTCAAGAAGGGGCCCCGGCAG 412

  134 AspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPr 150
      ||||||||||||||||||||||||||||||||||||||||||||:::||
  413 GATCCCAGCGCCATTGTGGAAAAGGCAGGGCTCAAGAAATCCATGGCCCC 462

  150 oThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleA 167
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  463 CACCCAGAACCGGGAGAAGTGGATCTACTACTGCAATGGCCACGGTATTG 512

  167 spIle.LeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIl 183
      ||||| |||||||||||||||||||||||||||||||||||||||||||
  513 ACATTCCTGAAGCTGGTAGCAGCCCAAGTGGGAAGCCAGTGGAAGGATAT 562

  183 eTyrGlnPheLeu 187
      |||||||||||||
  563 CTATCAGTTTCTG 575

seq_name: gb_est2:BG923601

seq_documentation_block:
LOCUS       BG923601      762 bp    mRNA            EST       05-JUN-2001
DEFINITION  602823467F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4952080 5',
            mRNA sequence.
ACCESSION   BG923601
VERSION     BG923601.1  GI:14304077
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10909  row: f  column: 17
            High quality sequence start: 6
            High quality sequence stop: 704.
FEATURES             Location/Qualifiers
     source          1..762
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4952080"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      203 a    213 c    196 g    150 t
ORIGIN


alignment_scores:
            Quality:  830.00            Length:    214
              Ratio:  4.278               Gaps:      9
 Percent Similarity:  90.654   Percent Identity: 84.579

alignment_block:
 US-09-733-956-2 x BG923601   ..

 Align seg 1/1  to: BG923601  from: 1  to: 762

    1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
      ||||||||||||:::||||||||||:::||||||||||||    ||||||
  136 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 185

   17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
       |||:::|||:::||||||||||||||||||||||||||||||:::    |
  186 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 235

   34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
      ||||||||    :::|||:::||||||||||||          |||:::|||
  236 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 285

   51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
      ||||||:::|||||||||||||||||||||||||||||||||     |||||
  286 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 335

   66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
      ||||||||||   |||||||||||||||||    ||||||:::|||||||
  336 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 382

   83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
      |||||||||||||||||||    ||||||||||||||||||||||||||
  383 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 432

  100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
      ||||||||||||||||||||||||||||||||||||||||||||:::|||||
  433 TTGCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGT 482

  116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  483 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 532

  133 lnAspProSerAlaIleValGlu.LysAlaGlyLeuLysLysSerMetTh 149
      ||||||||||||||||||||||| ||||||||||||||||||||||:::||
  533 AGGATCCCAGCGCCATAGTGGAACAAGGCGGGGCTGAAGAAGTCCCTGAC 582

  149 rProThrGlnAsnArgGluLysTrpIleTyr.TyrCysAsnGlyHisGly 165
      |||||||||||||||||||||||||||||||           ||||||||||||
  583 TCCCACCCAGAACCGGGAGAAATGGATCTACTAACCGCAACGGCCATGGT 632

  166 .IleAspIle.LeuLysLeuValAlaAla.GlnValGlySerGlnTrpLy 181
       ||||||||| ||||||||||||||||||| ||||||||||||||||||||
  633 AATTGACATCTTTGAAGCTTGTAGCAGCCCCAGGTGGGAAGCCAGTGGAA 682

  181 sAspIleTyrGlnPheLeuCysAsnAlaSerGluArg.GluValAla.Al 197
```

```
      |||||||||||||||||||||||||||||||||||| |||||||| ::
    683 GGACATCTATCAGTTTCTTTGCAACGCCAGCGAGAGGGGAGGTGGCGGAC 732

    197 aPheSerAsnGlyTyrThrAlaAspHis 206
        :||||||||||||||||||||||||||
    733 TTTCTCCAATGGATACACTGCAGATCAT 760

seq_name: gb_est1:BE478092

seq_documentation_block:
LOCUS       BE478092      521 bp    mRNA            EST       28-AUG-2000
DEFINITION  162175 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE478092
VERSION     BE478092.1  GI:9597625
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 521)
  AUTHORS   Sonstegard,T.S., Capuco,A.V., Van Tassell,C.P., Ashwell,M.S. and
            Wells,K.D.
  TITLE     Mapping of Expressed Sequence Tags from a normalized bovine mammary
            gland cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Sonstegard TS
            USDA, ARS, Beltsville Agricultural Research Center
            Bdlg. 200 Rm 2A, Beltsville, MD 20705, USA
            Tel: 301 504 8416
            Fax: 301 504 8414
            Email: tads@lpsi.barc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 5  row: P  column: 6
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .521
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="BARC 5BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled mRNA isolated from mammary
                     tissues at eight physiological, developmental, and disease
                     states."
BASE COUNT      135 a    167 c    135 g     84 t
ORIGIN


alignment_scores:
            Quality:  697.00            Length:    151
              Ratio:  4.840               Gaps:      0
 Percent Similarity:  95.364   Percent Identity: 90.728

alignment_block:
 US-09-733-956-2 x BE478092   ..

 Align seg 1/1  to: BE478092  from: 1  to: 521

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         :::|||||||||||||||||||||||||||||||||||||||||||||||
      68 ATCAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTGCC 117

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
          |||:::   ||||||||||||||||||||||||||||||||||||
     118 AAGTGGCACCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCGAGCA 167

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
           :::|||   |||||||||||||:::|||:::||||||||||:::|||||
     168 GGGAGGAGGGCGTGAACAAGACCTTCCCCAAGCTCCAGGTAGCGAACCAC 217

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     218 CAGCAAGGGCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 267

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||   |||||||
     268 GGCCACTGGGGGTGAGAAGTCAAGCACGCCCATCAAAGGCCTCAAGAGGG 317

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||   ||||||||||||||||||||||||||||||
     318 GCCACCCCAGGCAGAACCCACACAAGCATTTTGACATCAACGAGCACCTT 367

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     368 CCGTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTGGTGGTGATAGTGGT 417

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     418 GTGCAGTATCCGGAAAAGCTCGAGGACTCTCAAGAAGGGGCCCCGGCAGG 467

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||:::|||
     468 ATCCCAGCGCCATTGTGGAAAAGGCAGGGCTCAAGAAATCCATGGCCCCC 517

     151 Thr 151
         |||
     518 ACC 520

seq_name: gb_est1:AA072902

seq_documentation_block:
LOCUS       AA072902      398 bp    mRNA            EST       07-FEB-1997
DEFINITION  mm72f10.r1 Stratagene mouse macrophage (#937306) Mus musculus cDNA
            clone IMAGE:533995 5', mRNA sequence.
ACCESSION   AA072902
VERSION     AA072902.1  GI:1594632
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 398)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:320931
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 341.
FEATURES             Location/Qualifiers
     source          1. .398
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:533995"
                     /clone_lib="Stratagene mouse macrophage (#937306)"
                     /tissue_type="macrophage"
                     /dev_stage="WEHI-3 cell line"
```

```
                /lab_host="SOLR (kanamycin resistant)"
                /note="Organ: blood; Vector: pBluescript SK-; Site_1:
                EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;
                Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       95 a    118 c    111 g     74 t
ORIGIN


alignment_scores:
           Quality:  668.00             Length:    132
             Ratio:  5.099                Gaps:      0
 Percent Similarity:  99.242   Percent Identity: 95.455

alignment_block:
 US-09-733-956-2 x AA072902   ..

 Align seg 1/1  to: AA072902  from: 1  to: 398

     148 MetThrProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHi 164
         :::||||||||||||||||||||||||||||||||||||||   ||||||||
       1 CTGACTCCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCA 50

     164 sGlyIleAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpL 181
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      51 TGGTATTGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGA 100

     181 ysAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAla 197
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     101 AGGACATCTATCAGTTTCTTTGCAACGCCAGCGAGAGGGAGGTGGCGGCC 150

     198 PheSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGl 214
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TTCTCCAATGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCA 200

     214 nHisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerA 231
         |||||||||||||||||||||||||||||:::|||||||||||||||||||
     201 GCACTGGACCATCCGTGGCCCTGAGGCAGGCCTTGCCCAGCTCATTAGCG 250

     231 laLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeu 247
         ||||||||:::|||||||||||||||||||||||||||||||||||||||
     251 CCTTGCGCGAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTG 300

     248 MetGluAspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSe 264
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     301 ATGGAAGACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAG 350

     264 rProSerProLeuSerProSerProIleProSerProAsnAlaLys 279
         |||||||||||||||||||||||:::|||||||||||||:::|||
     351 CCCCAGTCCGCTTAGCCCGAGCCCCATGCCCAGTCCTAACGTGAAA 396

seq_name: gb_est2:BF129886

seq_documentation_block:
LOCUS       BF129886      692 bp    mRNA            EST       24-OCT-2000
DEFINITION  601817687F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:4041404 5',
            mRNA sequence.
ACCESSION   BF129886
VERSION     BF129886.1  GI:10968926
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 692)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM860  row: m  column: 21
            High quality sequence stop: 597.
FEATURES             Location/Qualifiers
     source          1. .692
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4041404"
                     /clone_lib="NIH_MGC_58"
                     /tissue_type="hypernephroma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.35 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      169 a    201 c    190 g    132 t
ORIGIN


alignment_scores:
           Quality:  660.50             Length:    220
             Ratio:  3.629                Gaps:     10
 Percent Similarity:  82.727   Percent Identity: 74.091

alignment_block:
 US-09-733-956-2 x BF129886   ..

 Align seg 1/1  to: BF129886  from: 1  to: 692

     165 GlyIleAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLy 181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      36 GGTATCGATATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAA 85

     181 sAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaP 198
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      86 AGATATCTATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTT 135

     198 heSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGln 214
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     136 TCTCCAATGGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAG 185

     215 HisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAl 231
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     186 CACTGGACCATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGC 235

     231 aLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuM 248
         |||||||||||||||||||||||    ||||||||||||||||||||||
     236 CCTGCGCCAGCACCGGAGAAACGATGT.GTGGAGAAGATTCGTGGGCTGA 284

     248 etGluAspThrThrGln.LeuGluThrAspLysLeuAlaLeuProMetSe 264
         |||||||||||||||| :::|||||||||||||||||||||||||||||||
     285 TGGAAGACACCACCCAGCGTGGAAACTGACAAACTAGCTCTCCCGATGAG 334

     264 rProSerProLeuSerProSerProIleProSer..ProAsnAlaLysLe 280
         |||||||||||||||||||||||||||||||||  ||||||||||||||
     335 CCCCAGCCCGCTTAGCCCGAGCCCCATCCCCAGCCACCAAACGCGAAACT 384

     280 uGluAsnSerAlaLeuLeuThrValGluProSerProGln..AspLysAs 296
         ||||||||||||||||||||||||||||      ||||||  ||||||||
     385 TGAGAATTCCGCTCTCCTGACGGTGGAGCTCTTCCCCCAGAGGACAAGAA 434

     296 nLysGlyPhePheValAspGluSer.GluProLeuLeuArgCysAspSer 312
```

```
    ||||||||||||||||||||||||| ||||||   :::   :::::::::
435 CAAGGGCTTCTTCGTGGATGAGTCGGGAGCCACCTGTCTCCGCTAGTGAC 484

313 Thr......SerSerGlySerSerAlaLeuSerArgAsn...GlySerPh 326
    :::      ::::::::::::|||||||||||||:::::   :::   ||
485 TCTCACATCCAGACGGCTGCCTCCGCGCTGAGCAAGGAACCGGCTCCTTT 534

326 eIleThr.LysGluLysLysAspThrValLeu.ArgGlnValArgLeuAs 342
    ||||||| |||||||||||||||||||::: |||||||||
535 TATTACCCAAAGAAAAGAAGGACACAGTTGTGGCGGCAGGTG........ 576

342 pProCysAspLeuGlnProIlePheAspAspMetLeuHis.......... 355
                        ::::::   |||::::::
577 ................ACCGCCTGGGAACCCATGGTGAATTGAAGTCTTA 610

356 .........................PheLeuAsnProGluGluLeuArg 363
                             ||||||   ||||||||||:::|||
611 TCTTTTGAATGACAATGGCTACCATATTCTTAATTCCTGAGGAGATGCGG 660
```

THIS PAGE BLANK (USPTO)

```
OM of: US-09-733-956-2 to: Issued_Patents_NA:*   out_format : pfs

Date: Jan 4, 2002  8:42 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09733956/runat_04012002_084025_15531/app_query.fasta_1.467
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09733956_@CGN1_1_80 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-733-956-2
Query length: 405
Database: Issued_Patents_NA:*
Database sequences: 351203
Database length: 113238999
Search time (sec): 88.950000



score_list:
Sequence            Strd Orig    ZScore    EScore Len  ! Documentation  ..
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-959-382-1 + 2091.00 4185.82  1.7e-225  2186  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-5 + 1914.00 3830.95  9.8e-206  1719  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-1 + 1914.00 3822.29  3.0e-205  3331  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-24 + 1527.00 3047.46  4.3e-162  1815  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-22 + 1527.00 3042.56  8.0e-162  2638  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-3 + 1269.00 2520.83  9.2e-133  2612  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-746-111-4 +  115.50  175.53  0.0397    6585  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-041-886-1 +  113.00  179.19  0.0249    3386  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-602-791-1 +  112.00  186.99  0.0091    1600  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-985-950-5 +  107.50  180.82  0.0202    1279  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-985-950-7 +  107.50  176.87  0.0334    1728  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-961-083-55 +  105.00  167.58  0.1102    2389  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-985-916-15 +  101.00  154.36  0.6005    3533  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-324-867-1 +   99.00  141.30  3.21      7032  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-989-925-2 +   96.50  153.06  0.7092    1947  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-208-887A-48 +   92.50  141.41  3.16      2556  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-428-517-1 -   92.50  102.21  481.96    50937  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-315-793-51 +   91.50  136.11  6.24      3282  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-229-515A-12 -   91.00  143.58  2.39      1717  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-645-865-12 -   91.00  143.58  2.39      1717  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-872-644-44 -   91.00  136.08  6.26      3044  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-297-494-44 -   91.00  136.08  6.26      3044  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-297-510-44 -   91.00  136.08  6.26      3044  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-479-532-44 -   91.00  136.08  6.26      3044  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-455-526-44 -   91.00  136.08  6.26      3044  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-455-525-44 -   91.00  136.08  6.26      3044  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-139-491-44 -   91.00  136.08  6.26      3044  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US92-03222-44 -   91.00  136.08  6.26      3044
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-928-941D-28 +   91.00  133.29  8.95      3767  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-928-941D-30 +   91.00  133.29  8.95      3767  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-280-590A-28 +   91.00  133.29  8.95      3767  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-280-590A-30 +   91.00  133.29  8.95      3767  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-459-568-3 +   90.00  127.09  19.83     5183  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-399-411-3 +   90.00  127.09  19.83     5183  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-516-859A-3 +   90.00  125.46  24.43     5868  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-2 +   89.50   37.73  6.5e+05   4403765
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-1 +   89.50   37.71  6.5e+05   4411529
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-1 -   89.50   37.71  6.5e+05   4411529
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-928-941D-2 +   89.00  132.66  9.71      2903  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-928-941D-3 +   89.00  132.66  9.71      2903  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-280-590A-2 +   89.00  132.66  9.71      2903
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-280-590A-3 +   89.00  132.66  9.71      2903
/cgn2_6/ptodata/2/ina/backfiles1.seq:5240838-4 -   89.00  130.88  12.20     3325  ! P
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-413-574-1 +   88.50  140.11  3.74      1522
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-418-782-1 -   88.50  135.07  7.12      2235

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-959-382-1

seq_documentation_block:
; Sequence 1, Application US/08959382
; Patent No. 6013476
;  GENERAL INFORMATION:
;    APPLICANT:  DEEN, KEITH
;    APPLICANT:  HURLE, MARK
;    APPLICANT:  YOUNG, PETER
;    APPLICANT:  TAN, K.B.
;    TITLE OF INVENTION:  Tumor Necrosis Related Receptor,
;    TITLE OF INVENTION:  TR7
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RATNER & PRESTIA
;      STREET:  P.O. BOX 980
;      CITY:  VALLEY FORGE
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/959,382
;      FILING DATE:  28-OCT-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/041,796
;      FILING DATE:  02-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PRESTIA, PAUL F
;      REGISTRATION NUMBER:  23,031
;      REFERENCE/DOCKET NUMBER:  GH-50017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-407-0700
;      TELEFAX:  610-407-0701
;      TELEX:  846169
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2186 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-959-382-1


alignment_scores:
            Quality: 2091.00              Length:    405
              Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x US-08-959-382-1   ..

 Align seg 1/1  to: US-08-959-382-1  from: 1  to: 2186

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     757 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 806

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     807 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 856
```

```
  34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 857 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 906

  51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 907 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 956

  67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 957 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1006

  84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1007 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1056

 101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1057 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1106

 117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1107 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1156

 134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1157 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1206

 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1207 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1256

 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1257 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1306

 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1307 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1356

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1357 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1406

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1407 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1456

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1457 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1506

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1507 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1556

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1557 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1606

 284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1607 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1656

 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1657 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1706

 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1707 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1756

 334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1757 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 1806

 351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1807 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 1856

 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1857 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 1906

 384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1907 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 1956

 401 LeuProAspLeuLeu 405
     ||||||||||||||||
1957 CTTCCTGACCTGCTG 1971
```

```
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-5

seq_documentation_block:
; Sequence 5, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1719
US-09-042-785A-5


alignment_scores:
            Quality: 1914.00              Length:    406
              Ratio:  4.883                Gaps:      2
 Percent Similarity:  96.552   Percent Identity: 92.857
```

```
alignment_block:
 US-09-733-956-2 x US-09-042-785A-5   ..

 Align seq 1/1  to: US-09-042-785A-5  from: 1  to: 1719

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||:::|||||||||:::||||||||||||   ||||||
     505 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 554

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         |||:::|||:::|||||||||||||||||||||||||:::   |
     555 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 604

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||   :::|||:::||||||||||||         |||:::|||
     605 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 654

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
         ||||||:::|||||||||||||||||||||||||||||||   |||||
     655 CAGCAAGCCCCCCACCACACACATTCTGAAGCTGCTGCCATCGTCCAT 704

      66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
         ||||||||||   ||||||||||||||||||   ||||||:::|||||||
     705 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 751

      83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
         ||||||||||||||||||   |||||||||||||||||||||||||||||
     752 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 801

     100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
     802 TTGCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGT 851

     116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     852 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 901

     133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
     902 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 951

     150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
         ||||||||||||||||||||||||||||||   ||||||||||||||||
     952 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 1001

     166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1002 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACA 1051

     183 leTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1052 TCTATCAGTTTCTTTGCAACGCCAGTGAGAGGGAGGTGGCGGCCTTCTCC 1101

     200 AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1102 AATGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCAGCACTG 1151

     216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1152 GACCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGC 1201

     233 rgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1202 GCCAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAA 1251

     250 AspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSe 266
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1252 GACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAGCCCCAG 1301

     266 rProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnS 283
         ||||||||||||||||||||||||||||||||:::|||||||||||||||
    1302 TCCGCTGAGCCCGAGCCCCATCCCCAGTCCTAACGTGAAACTTGAGAATT 1351

     283 erAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
         ||:::|||||||||||||||||||||||   ||||||||||||   |||
    1352 CCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAGTGCTTC 1401

     300 PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGl 316
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1402 TTCGTGGACGAGTCAGAGCCCCTTCTGCGTTGCGACTCCACATCCAGTGG 1451

     316 ySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysA 333
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1452 CTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAAGAAAAGAAGG 1501

     333 spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1502 ACACAGTGTTGCGGCAGGTCCGCCTGGACCCCTGTGACTTGCAGCCCATC 1551

     350 PheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGl 366
         ||||||||||||||||||:::||||||||||||||||||||||||||||||
    1552 TTTGATGACATGCTGCATATCCTGAACCCCGAGGAGCTGCGGGTGATTGA 1601

     366 uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleG 383
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1602 AGAGATTCCCCAGGCTGAGGACAAACTGGACCGCCTCTTCGAGATCATTG 1651

     383 lyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSer 399
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1652 GGGTCAAGAGCCAAGAAGCCAGCCAGACCCTCTTGGACTCTGTGTACAGT 1701

     400 HisLeuProAspLeuLeu 405
         ||||||||||||||||||
    1702 CATCTTCCTGACCTATTG 1719

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-1

seq_documentation_block:
; Sequence 1, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3331 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  344..2065
US-09-042-785A-1


alignment_scores:
           Quality: 1914.00            Length:    406
             Ratio:  4.883              Gaps:      2
 Percent Similarity:  96.552   Percent Identity: 92.857

alignment_block:
 US-09-733-956-2 x US-09-042-785A-1   ..

 Align seq 1/1  to: US-09-042-785A-1  from: 1  to: 3331

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||:::|||||||||:::|||||||||||   ||||||
     848 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 897

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
          |||:::|||:::|||||||||||||||||||||||||||:::   |
     898 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 947

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||   :::|||:::||||||||||||         |||:::|||
     948 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 997

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
         ||||||:::|||||||||||||||||||||||||||||||||   |||||
     998 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 1047

      66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
         |||||||||   ||||||||||||||||||   ||||||:::|||||||
    1048 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 1094

      83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
         ||||||||||||||||||   |||||||||||||||||||||||||||||
    1095 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 1144

     100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
    1145 TTGCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGT 1194

     116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1195 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 1244

     133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
         |||||||||||||||||||||||||||||||||||||||||||:::|||
    1245 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 1294

     150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
         ||||||||||||||||||||||||||||||   ||||||||||||||
    1295 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 1344

     166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1345 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACA 1394

     183 leTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1395 TCTATCAGTTTCTTTGCAACGCCAGTGAGAGGGAGGTGGCGGCCTTCTCC 1444

     200 AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1445 AATGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCAGCACTG 1494

     216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1495 GACCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGC 1544

      233 rgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1545 GCCAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAA 1594

      250 AspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSe 266
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1595 GACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAGCCCCAG 1644

      266 rProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnS 283
          ||||||||||||||||||||||||||||||||||:::|||||||||||||
     1645 TCCGCTGAGCCCGAGCCCCATCCCCAGTCCTAACGTGAAACTTGAGAATT 1694

      283 erAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
          ||:::|||||||||||||||||||||||   ||||||||||||   |||
     1695 CCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAGTGCTTC 1744

      300 PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGl 316
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1745 TTCGTGGACGAGTCAGAGCCCCTTCTGCGTTGCGACTCCACATCCAGTGG 1794

      316 ySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysA 333
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1795 CTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAAGAAAAGAAGG 1844

      333 spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1845 ACACAGTGTTGCGGCAGGTCCGCCTGGACCCCTGTGACTTGCAGCCCATC 1894

      350 PheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGl 366
          ||||||||||||||||||:::|||||||||||||||||||||||||||||
     1895 TTTGATGACATGCTGCATATCCTGAACCCCGAGGAGCTGCGGGTGATTGA 1944

      366 uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleG 383
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1945 AGAGATTCCCCAGGCTGAGGACAAACTGGACCGCCTCTTCGAGATCATTG 1994

      383 lyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSer 399
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1995 GGGTCAAGAGCCAAGAAGCCAGCCAGACCCTCTTGGACTCTGTGTACAGT 2044

      400 HisLeuProAspLeuLeu 405
          ||||||||||||||||||
     2045 CATCTTCCTGACCTATTG 2062

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-24

seq_documentation_block:
; Sequence 24, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  Mandragouras, Amy E
;       REGISTRATION NUMBER:  36,207
;       REFERENCE/DOCKET NUMBER:  MEI-001CP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617)227-7400
;       TELEFAX:  (617)742-4214
;   INFORMATION FOR SEQ ID NO:  24:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1815 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..1815
US-09-042-785A-24


alignment_scores:
            Quality: 1527.00              Length:    294
              Ratio:  5.194                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x US-09-042-785A-24

 Align seg 1/1  to: US-09-042-785A-24  from: 1  to: 1815

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 800

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     801 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 850

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     851 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 900

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     901 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 950

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     951 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1000

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1001 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1050

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1051 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1100

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1101 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1150

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1151 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1200

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1201 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1250

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1251 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1300

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1301 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1350

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1351 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1400

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1401 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1450

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1451 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1500

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1501 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1550

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1551 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1600

     284 laLeuLeuThrValGluProSerProGlnAsp 294
         |||||||||||||||||||||||||||||||
    1601 CTCTCCTGACGGTGGAGCCTTCCCCACAGGAT 1632

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-22

seq_documentation_block:
; Sequence 22, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2638 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  510..2327
```

```
US-09-042-785A-22


alignment_scores:
            Quality: 1527.00              Length:    294
              Ratio:  5.194                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x US-09-042-785A-22   ..

 Align seg 1/1  to: US-09-042-785A-22  from: 1  to: 2638

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1260 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1309

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1310 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1359

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1360 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1409

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1410 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1459

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1460 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1509

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1510 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1559

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1560 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1609

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1610 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1659

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1660 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1709

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1710 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1759

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1760 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1809

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1810 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1859

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1860 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1909

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1910 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1959

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1960 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 2009

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    2010 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 2059

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    2060 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 2109

     284 laLeuLeuThrValGluProSerProGlnAsp 294
         |||||||||||||||||||||||||||||||
    2110 CTCTCCTGACGGTGGAGCCTTCCCCACAGGAT 2141

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-042-785A-3

seq_documentation_block:
; Sequence 3, Application US/09042785A
; Patent No. 6194151
;  GENERAL INFORMATION:
;    APPLICANT:  Busfield, Samantha J
;    TITLE OF INVENTION:  NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,785A
;      FILING DATE:  17-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/938,896
;      FILING DATE:  26-SEP-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  MEI-001CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2612 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  190..951
US-09-042-785A-3


alignment_scores:
            Quality: 1269.00              Length:    259
              Ratio:  5.138                Gaps:      1
 Percent Similarity:  95.367   Percent Identity: 94.981

alignment_block:
 US-09-733-956-2 x US-09-042-785A-3   ..

 Align seg 1/1  to: US-09-042-785A-3  from: 1  to: 2612

      11 SerValArgProLysValLeuSerSerIleGlnGluGlyThrValProAs 27
         |||   |||||||||||||||||||||||||||||||||||||||||||||
     930 TCCCAAAGACCAAAGGTACTGAGTAGCATCCAGGAAGGGACAGTCCCTGA 979
```

```
     27 pAsnThrSerSerAlaArgGlyLysGluAspValAsnLysThrLeuProA 44
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    980 CAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGACCCTCCCAA 1029

     44 snLeuGlnValValAsnHisGlnGlnGlyProHisHisArgHisIleLeu 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1030 ACCTTCAGGTAGTCAACCACCAGCAAGGCCCCCACCACAGACACATCCTG 1079

     61 LysLeuLeuProSerMetGluAlaThrGlyGlyGluLysSerSerThrPr 77
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1080 AAGCTGCTGCCGTCCATGGAGGCCACTGGGGGCGAGAAGTCCAGCACGCC 1129

     77 oIleLysGlyProLysArgGlyHisProArgGlnAsnLeuHisLysHisP 94
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1130 CATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAGCATT 1179

     94 heAspIleAsnGluHisLeuProTrpMetIleValLeuPheLeuLeuLeu 110
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1180 TTGACATCAATGAGCATTTGCCCTGGATGATTGTGCTTTTCCTGCTGCTG 1229

    111 ValLeuValValIleValValCysSerIleArgLysSerSerArgThrLe 127
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1230 GTGCTTGTGGTGATTGTGGTGTGCAGTATCCGGAAAAGCTCGAGGACTCT 1279

    127 uLysLysGlyProArgGlnAspProSerAlaIleValGluLysAlaGlyL 144
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1280 GAAAAAGGGGCCCCGGCAGGATCCCAGTGCCATTGTGGAAAAGGCAGGGC 1329

    144 euLysLysSerMetThrProThrGlnAsnArgGluLysTrpIleTyrTyr 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1330 TGAAGAAATCCATGACTCCAACCCAGAACCGGGAGAAATGGATCTACTAC 1379

    161 CysAsnGlyHisGlyIleAspIleLeuLysLeuValAlaAlaGlnValGl 177
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1380 TGCAATGGCCATGGTATCGATATCCTGAAGCTTGTAGCAGCCCAAGTGGG 1429

    177 ySerGlnTrpLysAspIleTyrGlnPheLeuCysAsnAlaSerGluArgG 194
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1430 AAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAGAGGG 1479

    194 luValAlaAlaPheSerAsnGlyTyrThrAlaAspHisGluArgAlaTyr 210
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1480 AGGTTGCTGCTTTCTCCAATGGGTACACAGCCGACCACGAGCGGGCCTAC 1529

    211 AlaAlaLeuGlnHisTrpThrIleArgGlyProGluAlaSerLeuAlaGl 227
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1530 GCAGCTCTGCAGCACTGGACCATCCGGGGCCCCGAGGCCAGCCTCGCCCA 1579

    227 nLeuIleSerAlaLeuArgGlnHisArgArgAsnAspValValGluLysI 244
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   1580 GCTAATTAGCGCCCTGCGCCAGCACCGGAGAAACGATGTTGTGGAGAAGA 1629

    244 leArgGlyLeuMetGluAspThrThrGln..................... 253
        ||||||||||||||||||||||||||||
   1630 TTCGTGGGCTGATGGAAGACACCACCCAGGTAATGGAGCCCTTGTTGTGT 1679

    254 ...LeuGluThrAspLysLeuAlaLeu 261
           :::   ||||||   ||| ||| 
   1680 GTCATTACCACCGACCTATTGCCCCTA 1706

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-746-111-4

seq_documentation_block:
; Sequence 4, Application US/08746111
; Patent No. 6066778
;  GENERAL INFORMATION:
;    APPLICANT:  Ginsburg, David
;    APPLICANT:  Cui, Jisong
;    TITLE OF INVENTION:  Compositions And Methods For Screening
;    TITLE OF INVENTION:  Compounds For Anticoagulant Activity
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/746,111
;      FILING DATE:  06-NOV-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  UM-02536
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6585 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA"
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..6554
US-08-746-111-4


alignment_scores:
            Quality:  115.50              Length:    488
              Ratio:  0.516                 Gaps:     20
 Percent Similarity:  45.902   Percent Identity: 19.877

alignment_block:
 US-09-733-956-2 x US-08-746-111-4   ..

 Align seq 1/1  to: US-08-746-111-4  from: 1  to: 6585

       5 GluSerAsnSerSerAlaSerValArgProLysValLeuSerSerIleGl 21
         :::|||||||||||||              ::::::||||||:::|||
    2322 GACTCAAACTCTTCA..............CGAATCCTTAGTAAAATCAT 2356

      21 nGluGlyThrValProAsp......AsnThrSerSerAlaArgGlyLysG 36
          ::::::::::::   |||       :::|||      :::   |||
    2357 CAATAATAACCTCAAAGACTTTCAAAGAACACTTCCTGGCTCAGGAGCCA 2406

      36 luAspValAsnLysThrLeuProAsnLeu..................... 45
           ::::::      |||   ||||||
    2407 CCGTGGCTGGTACCCTCCTTAGAAACCTCATTGGCTTAGATGAGAACTTC 2456

      46 .........GlnValValAsnHisGlnGlnGlyProHisHisArgHis.. 58
                 :::   ::::::|||:::::::::   :::|||::::::
    2457 GTCCTCAACTCTTCTACAGAACATCGTTCCAGCTCATATCATGAAAATGA 2506

      59 ............................IleLeuLysLeuLeuProSerM 66
                                     ::::::   |||||||||
    2507 TATGGAAAATCCACAGTCAAACATCACAATGGTATACCTACTTCCTCTTG 2556

      66 etGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLys 82
                   :::|||:::::::::::          |||      |||
    2557 GTCCAAAAGGATCTGGGAATCGAGAACAAGATAAACCTAAAACCATCAAG 2606

      83 ArgGlyHisProArgGlnAsnLeuHisLysHisPhe.............. 94
           |||:::|||::::::      |||:::
    2607 ACAGGAAGACCCCACATGATGAAGCACAGGTTCTCCTGGATGAAAGCGCC 2656
```

```
 95 .......AspIleAsnGluHisLeuProTrpMetIleValLeuPheLeuL 109
                  :::::::|||
2657 AGCTGGTAAAACTGGGAGGCAT........................... 2678

109 euLeuValLeuValValIleValValCysSerIleArgLysSerSerArg 125
                                 |||      |||:::|||
2679 ..........................TCAAACCCAAAGAATTCGTAT 2699

126 ThrLeuLysLysGlyProArgGlnAspProSerAlaIleValGluLysAl 142
    :::      |||:::   ::::::   ||||||   ::::::
2700 TCTGGAATGAAGTCTGAGGAGGACATTCCTAGCGAGTTGATACCC..... 2744

142 aGlyLeuLysLysSerMetThrProThrGlnAsnArgGluLysTrp.... 157
      ||||||:::::::::|||           :::::::::|||
2745 ....TTAAAGCAAAAGATCACTTCCAAATTTCTGAATAGACGATGGCGTG 2790

158 .......................IleTyrTyrCysAsnGlyHisGly 165
                            |||      :::||||||:::
2791 TGGCTTCTGAAAAGGGTAGTTATGAAATAATAGCAGCAAATGGTGAAGAC 2840

166 IleAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAs 182
      |||:::   ||||||:::
2841 ACAGATGTGGATAAGCTGACC........................... 2861

182 pIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheS 199
                      |||:::   :::::::::     :::
2862 ..................AACAGTCCTCAAAATCAGAATATCACAGTAC 2892

199 erAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHis 215
    :::|||   :::::::::|||   :::
2893 CTCGGGGAGAGAGCACCTCTCACACAAAC..................... 2921

216 TrpThrIleArgGlyProGlu.............AlaSerLeuAlaGl 227
      |||   |||   |||:::               :::::::::::::
2922 ...ACAACAAGAAAGCCAAGTGACCTCCCAACATTTTCTGGAGTTGGACA 2968

227 nLeuIleSerAlaLeuArgGlnHisArgArgAsnAspValValGluLysI 244
    :       :::||||||:::::::::|||:::      :::|||
2969 TAAATCTCCACATGTAAGACAGGAGGAAGAAAACAGTGGTTTTCAGAAAA 3018

244 leArgGlyLeuMetGluAspThrThrGlnLeuGluThrAspLysLeuAla 260
      :::   :::::::::        :::   ::::::   ||||||||
3019 GACAGTTATTCATCAGGACACGGAAGAAGAAGAAAAATAAGAAGCTTGCA 3068

261 Leu......ProMetSerProSerProLeuSerPro.............Se 271
    |||      |||:::||||||      :::::::|||           ::
3069 CTACACAGTCCTCTATCTCCAAGGGGCTTTGACCCTTTGAGAGGACATAA 3118

271 rProIleProSerProAsnAlaLysLeuGluAsnSerAlaLeuLeu.... 286
    :     |||   |||:::   :::|||   |||   :::||||||
3119 CCATTCCCCATTTCCAGACAGGAGACTACTTAATCACTCACTGTTACTCC 3168

287 .................ThrValGluProSerProGlnAspLysAsnLys 297
                    :::::::::|||:::   ::::::   :::
3169 ACAAGTCCAATGAAACAGCTCTTTCTCCAGACCTGAACCAGACCTCTCCT 3218

298 GlyPhePheValAspGluSerGluPro.........LeuLeuArgCysAs 311
    ::::::   :::|||:::|||   |||                    ::
3219 TCAATGAGTACGGACAGGTCACTTCCTGACTATAATCAGTACTCGAAAAA 3268

311 pSerThrSerSerGlySerSerAlaLeuSerArgAsnGlySerPheIleT 328
    ::::|||::::::   ||||||:::|||:::         |||      :
3269 TGACACTGAGCAGATGAGCTCTTCTTTAGATCTTTATCAGTCAGTGCCCG 3318

328 hrLysGluLysLysAspThrValLeuArgGln.................. 338
    :::::|||   :::   |||          |||
3319 CAGAGGAACACTCTCCAACATTTCCTGCCCAAGATCCTGATCAAACACAC 3368

339 .................ValArgLeuAspProCysAspLeuGlnProIl 349
                       |||   :::|||   :::||||:::
3369 TCTACCACAGATCCTAGCTACAGATCCTCTCCGCCAGAGCTCAGCCAGGG 3418

349 ePheAsp...AspMetLeuHisPheLeuAsnProGluGluLeuArgValI 365
      :::|||   |||:::   |||   :::   |||::::::::::   :::
3419 GCTTGATTATGACCTAAGTCATGACTTTTACCCTGATGACATTGGTCTAA 3468

365 leGluGluIleProGlnAlaGluAspLysLeuAsp.............. 376
      :::   :::|||:::   :::::::|||   :::
3469 CATCTTTCTTTCCAGACCAAAGTCAAAAGTCATCTTTCTCTTCAGATGAT 3518

377 ........................ArgLeuPheGluIleIleGlyVa 384
                                        ||||||   |||
3519 GACCAAGCAATCCCTTCCTCAGACTTAAGCCTCTTTACCATC........ 3560

384 lLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHisL 401
      |||   |||   :::|||||||:::         :::|||     |
3561 ....TCTCCAGAATTGGATCAGACAATT.........ATTTACCCAGACC 3597

401 euProAspLeuLeu 405
    ||   :::||||||
3598 TGGATCAGTTGCTC 3611
```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-041-886-1

```
seq_documentation_block:
; Sequence 1, Application US/09041886
; Patent No. 6235872
;  GENERAL INFORMATION:
;    APPLICANT:  Bredesen, Dale E.
;    APPLICANT:  Rabizadeh, Sharroz
;    TITLE OF INVENTION:  Proapoptotic Peptides, Dependence
;    TITLE OF INVENTION:  Polypeptides and Methods of Use
;    NUMBER OF SEQUENCES:  72
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/041,886
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LJ 2626
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3386 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  114..1395
US-09-041-886-1
```

```
alignment_scores:
            Quality:  113.00              Length:    300
              Ratio:  0.837                 Gaps:     12
 Percent Similarity:  45.000   Percent Identity: 22.333

alignment_block:
```

```
US-09-733-956-2 x US-09-041-886-1   ..

Align seg 1/1  to: US-09-041-886-1  from: 1  to: 3386

     62 LeuLeuProSerMetGluAlaThrGlyGlyGluLysSerSerThrProIl 78
        |||   |||:::      :::   ||||||      :::|||
    910 TTGTGGCCTACATAGCCTTCAAGAGGTGGAACAGCTGCAAGCAGAACAAG 959

     78 eLysGlyPro.................LysArgGlyHisProArgGlnA 89
        |||   |||                  :::|||   |||   |||:::
    960 CAAGGAGCCAACAGCCGGCCAGTGAACCAGACGCCCCCACCAGAGGGAG. 1008

     89 snLeuHisLysHisPheAspIleAsnGluHisLeuProTrpMetIleVal 105
              |||:::         :::         ||||||
   1009 ........AAAAACTCCACAGCGACAGTGGCATCTCCGTGG......... 1041

    106 LeuPheLeuLeuLeuValLeuValValIleValValCysSerIleArgLy 122
                                    .  ::::::
   1042 ..............................ACAGCCAGAGCCTGCAT 1058

    122 sSerSerArgThrLeuLysLysGlyProArgGlnAspProSerAlaIleV 139
        :::|||        ::::::   ||||||      ||||||   :::
   1059 GACCAGCAGCCCCACACGCAGACAGCCTCGGGCCAGGCCCTCAAGGGTGA 1107

    139 alGluLysAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgGlu 155
              :::|||||||   :::|||:::   |||:::::::::::|||
   1108 C.......GGAGGCCTCTACAGCAGCCTGCCCCCAGCCAAGCGGGAGGAG 1151

    156 LysTrpIleTyrTyrCysAsnGlyHisGlyIleAspIleLeuLysLeuVa 172
                                          :::   ||||||::
   1152 ................................GTGGAGAAGCTTCT 1165

    172 lAlaAlaGlnValGlySerGlnTrpLysAspIleTyrGlnPheLeuCysA 189
        :   :::::::::|||:::   |||:::   :::
   1166 CAACGGCTCTGCGGGGGACACCTGGCGGCACCTG.............. 1199

    189 snAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyrThrAlaAsp 205
                            |||   :::   ||||||      :::
   1200 ....................GCGGGCGAGCTGGGCTACCAGCCCGAG 1226

    206 HisGluArgAlaTyr.....................AlaAlaLeuGl 214
        |||      ::::::                        |||   |||
   1227 CACATAGACTCCTTTACCCATGAGGCCTGCCCCGTTCGCGCCCTGCTTGC 1276

    214 nHisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerA 231
           |||:::   :::      :::|||:::|||      |||:::::::|
   1277 AAGCTGGGCCACCCAG...GACAGCGCCACACTGGACGCCCTCCTGGCCG 1323

    231 laLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeu 247
        ||||||||:::   :::|||  ||||:::||||||         :::|||
   1324 CCCTGCGCCGCATCCAGCGAGCCGACCTCGTGGAG.........AGTCTG 1364

    248 MetGluAspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSe 264
           :::::::::|||      :::                  |||:::
   1365 TGCAGTGAGTCCACTGCCACATCC.................CCGGTGTG 1396

    264 rProSerProLeuSerProSerProIleProSerProAsnAlaLysLeuG 281
        :::|||   ||||||   |||   |||      :::   :::   :
   1397 AGCCCAACCGGGGAGCCCCCGCCCCGCCCCACATTCCGACAACCGATGCT 1446

    281 luAsnSerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLys 297
        ::::::::   |||   :::   :::|||   |||             :::
   1447 CCAGCCAACCCCTGTGGAGCCCGCACCCCCACCCTTTGGGGGGGGCCCGC 1496

    298 GlyPhePheValAspGluSerGluProLeuLeuArgCysAspSerThrSe 314
              :::::::::|||         ||||||   ::::::   ::
   1497 CTGGCAGAACTGAGCTCCTCTGGGCAGGACCTCAGAGTCCAGGCCCCAAA 1546

    314 rSerGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluL 331
        :              : |||      |||   |||   ||:::::::
   1547 ACCACAGCCCTGTCAGTGCAGCCCGTGTGGCCCCTTC...ACTTCTGACC 1593

    331 ysLysAspThrValLeuArgGlnValArgLeu.........AspProCys 344
```

```
              :::         |||:::|||   |||         :::|||||||
   1594 ACACTTCCTGTCCAGAGAGAGAAGTGCCCCTGCTGCCTCCCCAACCCTGC 1643

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-602-791-1

seq_documentation_block:
; Sequence 1, Application US/08602791
; Patent No. 6074836
;  GENERAL INFORMATION:
;    APPLICANT:  BORDIGNON, Claudio
;    APPLICANT:  MAVILIO, Fulvio
;    TITLE OF INVENTION:  METHOD OF MARKING EUKARYOTIC CELLS
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 Fifteenth Street, N. W.,
;      STREET:  Suite 330 - G Street Lobby
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  2005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/602,791
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IT RM93/A000587
;      FILING DATE:  01-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO 95/06723
;      FILING DATE:  11-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLESNER, Sharon N.
;      REGISTRATION NUMBER:  36,335
;      REFERENCE/DOCKET NUMBER:  P1614-4007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1600 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  one-of(943)
;      OTHER INFORMATION:  /function= "PvuII cleavage site"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  one-of(1512)
;      OTHER INFORMATION:  /function= "SstI cleavage site"
US-08-602-791-1


alignment_scores:
           Quality:  112.00              Length:    245
             Ratio:  1.009                 Gaps:     10
 Percent Similarity:  45.306   Percent Identity: 22.449

alignment_block:
 US-09-733-956-2 x US-08-602-791-1   ..

 Align seg 1/1  to: US-08-602-791-1  from: 1  to: 1600

      62 LeuLeuProSerMetGluAlaThrGlyGlyGluLysSerSerThrProIl 78
         |||   |||:::      :::   ||||||      :::|||
```

```
    910 TTGTGGCCTACATAGCCTTCAAGAGGTGGAACAGCTGCAAGCAGAACAAG 959

     78 eLysGlyPro.................LysArgGlyHisProArgGlnA 89
        |||   |||                 :::|||   |||   |||:::
    960 CAAGGAGCCAACAGCCGGCCAGTGAACCAGACGCCCCCACCAGAGGGAG. 1008

     89 snLeuHisLysHisPheAspIleAsnGluHisLeuProTrpMetIleVal 105
              |||:::       :::       ||||||
   1009 ........AAAAACTCCACAGCGACAGTGGCATCTCCGTGG......... 1041

     106 LeuPheLeuLeuLeuValLeuValValIleValValCysSerIleArgLy 122
                                      ::::::
    1042 ..............................ACAGCCAGAGCCTGCAT 1058

     122 sSerSerArgThrLeuLysLysGlyProArgGlnAspProSerAlaIleV 139
         :::|||       ::::::   ||||||     ||||||   :::
    1059 GACCAGCAGCCCCACACGCAGACAGCCTCGGGCCAGGCCCTCAAGGGTGA 1107

     139 alGluLysAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgGlu 155
              :::||||||   :::|||:::   |||:::::::::::::|||
    1108 C.......GGAGGCCTCTACAGCAGCCTGCCCCCAGCCAAGCGGGAGGAG 1151

     156 LysTrpIleTyrTyrCysAsnGlyHisGlyIleAspIleLeuLysLeuVa 172
                                         :::   ||||||::
    1152 ..............................GTGGAGAAGCTTCT 1165

     172 lAlaAlaGlnValGlySerGlnTrpLysAspIleTyrGlnPheLeuCysA 189
         :   ::::::::::|||:::   |||:::   :::
    1166 CAACGGCTCTGCGGGGGACACCTGGCGGCACCTG.............. 1199

     189 snAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyrThrAlaAsp 205
                          |||   :::   ||||||      :::
    1200 ..................GCGGGCGAGCTGGGCTACCAGCCCGAG 1226

     206 HisGluArgAlaTyr.....................AlaAlaLeuGl 214
         |||      ::::::                    |||   |||
    1227 CACATAGACTCCTTTACCCATGAGGCCTGCCCCGTTCGCGCCCTGCTTGC 1276

     214 nHisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerA 231
             |||:::   :::      :::|||:::|||      |||::::::|
    1277 AAGCTGGGCCACCCAG...GACAGCGCCACACTGGACGCCCTCCTGGCCG 1323

     231 laLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeu 247
         ||||||||:::   :::|||   |||:::||||||          :::|||
    1324 CCCTGCGCCGCATCCAGCGAGCCGACCTCGTGGAG.........AGTCTG 1364

     248 MetGluAspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSe 264
            ::::::::::|||      :::                  |||:::
    1365 TGCAGTGAGTCCACTGCCACATCC.................CCGGTGTG 1396

     264 rProSerProLeuSerProSerProIleProSerProAsnAlaLysLeuG 281
         :::|||   ||||||   |||   |||      :::   :::   :
    1397 AGCCCAACCGGGGAGCCCCCGCCCCGCCCCACATTCCGACAACCGATGCT 1446

     281 luAsnSerAlaLeuLeuThrValGluProSerPro 292
         ::::::::   |||   :::   :::|||   |||
    1447 CCAGCCAACCCCTGTGGAGCCCGCACCCCCACCCT 1481

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-985-950-5

seq_documentation_block:
; Sequence 5, Application US/08985950
; Patent No. 6140076
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1279 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  155..1015
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1247
;      OTHER INFORMATION:  /note= "nucleotide 1247 designated
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  218..1015
US-08-985-950-5


alignment_scores:
            Quality:  107.50              Length:    338
              Ratio:  0.773                 Gaps:     17
 Percent Similarity:  41.124   Percent Identity: 20.414

alignment_block:
 US-09-733-956-2 x US-08-985-950-5   ..

 Align seg 1/1  to: US-08-985-950-5  from: 1  to: 1279

      27 AspAsnThrSerSerAlaArgGlyLysGluAspValAsnLysThrLeuPr 43
         ::::::   |||:::   :::      |||||||||:::::::::   ||
     341 GAGAGTAGATCCACATACAATGATACTGAAGATGTGTCTCAAGCTAGTCC 390

      43 oAsnLeu.....................GlnValValAsnHisGlnG 52
         |:::                      :::|||   :::   :::
     391 ATCTGAGTCAGAGGCCAGATTCCGCATTGACTCAGTAAGTGAAGGAAATG 440

      52 lnGlyProHis.................................His 56
           ||||||:::                                 :::
     441 CCGGGCCTTATCGCTGCATCTATTATAAGCCCCCTAAATGGTCTGAGCAG 490

      57 ArgHisIleLeuLysLeuLeuProSerMetGluAlaThrGlyGlyGluLy 73
                 |||:::||||||   :::   |||:::::||||||
     491 AGTGACTACCTGGAGCTGCTGGTGAAA...GAAACCTCTGGAGGCCCGGA 537

      73 sSerSerThrProIleLysGlyProLysArgGly.........HisProA 87
```

```
     538 CTCCCCGGACACAGAGCCCGGCTCCTCAGCTGGACCCACGCAGAGGCCGT 587

      87 rgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeuPro..... 101
     588 CGGACAACAGTCAC.............AATGAGCATGCACCTGCTTCC 622

     102 .................................TrpMetIleValLe 106
     623 CAAGGCCTGAAAGCTGAGCATCTGTATATTCTCATCGGGGTCTCAGTGGT 672

     106 uPheLeuLeuLeuValLeuValValIleValValCysSerIleArgLysS 123
     673 CTTCCTCTTCTGTCTCCTCCTCCTGGTCCTCTTCTGCCTCCATCGCCAGA 722

     123 erSerArgThrLeuLysLysGlyProArgGlnAspProSerAlaIleVal 139
     723 ATCAG......ATAAAGCAGGGGCCCCCCAGAAGCAAGGAC......... 757

     140 GluLysAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgGluLy 156
     758 ..................GAGGAGCAGAAGCCACAGCAGAGGCCTGACCT 789

     156 sTrpIleTyrTyrCysAsnGlyHisGlyIleAspIleLeuLysLeuValA 173
     790 G.......................GCTGTTGATGTTCTAGAGAGGACAG 815

     173 laAlaGlnVal.........GlySerGlnTrpLysAspIleTyrGlnPhe 186
     816 CAGACAAGGCCACAGTCAATGGACTTCCTGAGAAGGAC............ 853

     187 LeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyrTh 203
     854 ............AGAGAGACGGACACCTCGGCCCTGGCTGCAGGG..... 886

     203 rAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpThrIleArgG 220
     887 .AGTTCCCAGGAGGTGACGTATGCTCAGCTGGACCACTGGGCCCTC.... 931

     220 lyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgGlnHisArg 236

     931 ..................................................... 931

     237 ArgAsnAspValValGluLysIleArgGlyLeuMetGluAspThrThrGl 253

     931 ..................................................... 931

     253 nLeuGluThrAspLysLeuAlaLeuProMetSerPro.SerProLeuSer 269
     932 .......ACACAGAGGACAGCCCGGGCTGTGTCCCCACAGTCCACAAAGC 974

     270 ProSerProIleProSerProAsnAlaLysLeuGluAsnSerAlaLeuLe 286
     975 CCATGGCCGAGTCCATCA.............................. 992

     286 uThrValGluProSerProGlnAspLysAsnGlyPhePheValAspG 303
     993 .CGTATGCAGCCGTTGCCA............................G 1011

     303 luSerGluProLeu.................LeuArgCysAspSerThr 313
    1012 ACACTGACCCCATACCCACCTGGCCTCTGCACCTGAGGGTAGAAAGTCAC 1061

     314 SerSerGlySer 317
    1062 TCTAGGAAAAGC 1073

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-985-950-7

seq_documentation_block:
; Sequence 7, Application US/08985950
; Patent No. 6140076
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1728 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  69..929
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  132..929
US-08-985-950-7


alignment_scores:
            Quality:  107.50            Length:    338
              Ratio:  0.773              Gaps:     17
 Percent Similarity:  41.124   Percent Identity: 20.414

alignment_block:
 US-09-733-956-2 x US-08-985-950-7   ..

 Align seg 1/1  to: US-08-985-950-7  from: 1  to: 1728

      27 AspAsnThrSerSerAlaArgGlyLysGluAspValAsnLysThrLeuPr 43
     255 GAGAGTAGATCCACATACAATGATACTGAAGATGTGTCTCAAGCTAGTCC 304

      43 oAsnLeu.....................GlnValValAsnHisGlnG 52
     305 ATCTGAGTCAGAGGCCAGATTCCGCATTGACTCAGTAAGTGAAGGAAATG 354

      52 lnGlyProHis.................................His 56
     355 CCGGGCCTTATCGCTGCATCTATTATAAGCCCCCTAAATGGTCTGAGCAG 404
```

```
 57 ArgHisIleLeuLysLeuLeuProSerMetGluAlaThrGlyGlyGluLy 73
405 AGTGACTACCTGGAGCTGCTGGTGAAA...GAAACCTCTGGAGGCCCGGA 451

 73 sSerSerThrProIleLysGlyProLysArgGly.........HisProA 87
452 CTCCCCGGACACAGAGCCCGGCTCCTCAGCTGGACCCACGCAGAGGCCGT 501

 87 rgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeuPro...... 101
502 CGGACAACAGTCAC..............AATGAGCATGCACCTGCTTCC 536

102 ..................................TrpMetIleValLe 106
537 CAAGGCCTGAAAGCTGAGCATCTGTATATTCTCATCGGGGTCTCAGTGGT 586

106 uPheLeuLeuLeuValLeuValValIleValValCysSerIleArgLysS 123
587 CTTCCTCTTCTGTCTCCTCCTCCTGGTCCTCTTCTGCCTCCATCGCCAGA 636

123 erSerArgThrLeuLysLysGlyProArgGlnAspProSerAlaIleVal 139
637 ATCAG......ATAAAGCAGGGGCCCCCCAGAAGCAAGGAC......... 671

140 GluLysAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgGluLy 156
672 ..................GAGGAGCAGAAGCCACAGCAGAGGCCTGACCT 703

156 sTrpIleTyrTyrCysAsnGlyHisGlyIleAspIleLeuLysLeuValA 173
704 G.......................GCTGTTGATGTTCTAGAGAGGACAG 729

173 laAlaGlnVal.........GlySerGlnTrpLysAspIleTyrGlnPhe 186
730 CAGACAAGGCCACAGTCAATGGACTTCCTGAGAAGGAC............ 767

187 LeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyrTh 203
768 ............AGAGAGACGGACACCTCGGCCCTGGCTGCAGGG..... 800

203 rAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpThrIleArgG 220
801 .AGTTCCCAGGAGGTGACGTATGCTCAGCTGGACCACTGGGCCCTC.... 845

220 lyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgGlnHisArg 236
845 ..................................................... 845

237 ArgAsnAspValValGluLysIleArgGlyLeuMetGluAspThrThrGl 253
845 ..................................................... 845

253 nLeuGluThrAspLysLeuAlaLeuProMetSerPro.SerProLeuSer 269
846 .......ACACAGAGGACAGCCCGGGCTGTGTCCCCACAGTCCACAAAGC 888

270 ProSerProIleProSerProAsnAlaLysLeuGluAsnSerAlaLeuLe 286
889 CCATGGCCGAGTCCATCA.............................. 906

286 uThrValGluProSerProGlnAspLysAsnLysGlyPhePheValAspG 303
907 .CGTATGCAGCCGTTGCCA............................G 925

303 luSerGluProLeu.................LeuArgCysAspSerThr 313
926 ACACTGACCCCATACCCACCTGGCCTCTGCACCTGAGGGTAGAAAGTCAC 975

314 SerSerGlySer 317
976 TCTAGGAAAAGC 987

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-961-083-55
```

```
seq_documentation_block:
; Sequence 55, Application US/08961083
; Patent No. 6159469
;  GENERAL INFORMATION:
;    APPLICANT:  Choi et. al.
;    TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;    NUMBER OF SEQUENCES:  452
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,083
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2389 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-083-55


alignment_scores:
            Quality:  105.00              Length:    534
              Ratio:  0.482                 Gaps:     23
 Percent Similarity:  40.824   Percent Identity: 17.978

alignment_block:
 US-09-733-956-2 x US-08-961-083-55   ..

 Align seg 1/1  to: US-08-961-083-55  from: 1  to: 2389

       5 GluSerAsnSerSerAlaSerValArgProLysValLeuSerSerIleGl 21
      41 GAAAATAATCGTGTTTCCTATATAGATGGAAAACAAGCGACGCAAAAAAC 90

      21 nGluGlyThrValProAspAsnThrSerSerAlaArgGly......... 34
      91 GGAGAATTTGACTCCTGATGAGGTTAGCAAGCGTGAAGGAATCAATGCTG 140

      34 ..................................................... 34

     141 AGCAAATCGTCATCAAGATAACAGACCAAGGCTATGTCACTTCACATGGC 190

      34 ..................................................... 34

     191 GACCACTATCATTATTACAATGGTAAGGTTCCTTATGACGCTATCATCAG 240

      35 .............LysGluAspValAsnLysThrLeuProAsnLeuGlnV 47
     241 TGAAGAATTACTCATGAAAGATCCAAACTATAAGCTAAAAGATGAGGATA 290
```

```
 47 alValAsnHisGlnGlnGlyProHisHisArgHisIleLeuLysLeu... 62
    ::|||||||:::   :::|||           ::::::::::|||:::
291 TTGTTAATGAGGTCAAGGGT.........GGATATGTTATCAAGGTAGAT 331

 63 ..................LeuProSerMetGluAlaThrGlyGlyGluLy 73
                       |||   :::         :::   :::   ::
332 GGAAAATACTATGTTTACCTTAAGGATGCTGCCCACGCGGATAACGTCCG 381

 73 sSerSerThrProIleLysGlyProLysArgGlyHisProArgGlnAsnL 90
    :::::::      |||:::      |||:::   |||   :::
382 TACAAAAGAGGAAATCAATCGACAAAAACAAGAGCATAGTCAA....... 424

 90 euHisLysHis.PheAspIleAsnGluHis.........LeuProTrp.. 102
      |||::::::    :::::::::|||            |||||||||
425 ..CATCGTGAAGGTGGAACTCCAAGAAACGATGGTGCTGTTGCCTTGGCA 472

103 ............................MetIleValLeuPheLeuLeuL 110
                                |||::::::   ::::::||||
473 CGTTCGCAAGGACGCTATACTACAGATGATGGTTATATCTTTAATGCTTC 522

110 euVal................................LeuValValIle 115
    ||:::                               |||:::   |||
523 TGATATCATAGAGGATACTGGTGATGCTTATATCGTTCCTCATGGAGATC 572

116 ValValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProAr 132
    ::::::   ::::::   :::   |||       |||   :::
573 ATTACCATTACATTCCTAAGAATGAGTTATCAGCTAGCGAGTTGGCTGCT 622

132 gGlnAspProSerAlaIleValGlu.LysAlaGlyLeuLysLysSerMet 148
     |||   ||||||   :::|||||| |||:::         |||:::::::
623 GCAGAAGCCTTCCTATCTGGTCGAGGAAATCTGTCAAATTCAAGAACCTA 672

149 ThrProThrGlnAsnArg............GluLys.TrpIle....... 158
    :::|||||||:::   |||           :::||| |||:::
673 TCGCCGACAAAATAGCGATAACACTTCAAGAACAAACTGGGTACCTTCTG 722

159 ........................TyrTyrCysAsnGlyHisGlyIle 166
                              |||:::   :::
723 TAAGCAATCCAGGAACTACAAATACTAACACAAGCAACAACAGCAACACT 772

167 AspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIl 183
    :::                 :::   ::::::   |||      |||:::::
773 AACAGTCAAGCAAGTCAAAGTAATGACATTGATAGTCTCTTGAAACAGCT 822

183 eTyrGlnPheLeuCysAsnAlaSerGluArgGluVal............. 195
    :|||::::::          |||:::|||:::|||
823 CTACAAACTG......CCTTTGAGTCAACGACATGTAGAATCTGATGGCC 866

196 .......................AlaAlaPheSerAsnGlyTyrThr 203
                           :::      ::::::|||   :::
867 TTGTCTTTGATCCAGCACAAATCACAAGTCGAACAGCTAGAGGTGTTGCA 916

204 AlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpThrIleArgGl 220
    :::   |||         |||      :::   :::      :::
917 GTGCCACACGGAGATCATTACCACTTCATCCCTTACTCTCAAATGTCTGA 966

220 yProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgGlnHisArgA 237
      |||      :::|||::::::|||             ::::::|||
967 ATTGGAAGAACGAATCGCTCGTATTATTCCCCTT......CGTTATCGTT 1010

237 rgAsnAspValValGluLysIleArgGlyLeuMetGluAspThrThrGln 253
     |||      |||         |||
1011 CAAACCATTGGGTACCAGATTCAAGG..................... 1036

254 LeuGluThrAspLysLeuAlaLeuProMetSerProSerProLeu..... 268
                          |||   :::|||||||||
1037 ......................CCAGAACAACCAAGTCCACAACCGAC 1062

269 .........SerProSerProIleProSerProAsnAlaLysLeuGluA 282
            ||||||:::|||   |||:::||||||   |||:::::::
1063 TCCGGAACCTAGTCCAGGCCCGCAACCTGCACCAAATCTTAAAATAGACT 1112

282 snSerAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGly 298
```

```
       ::::::::   |||   |||:::         :::       ::::::|||
1113 CAAATTCTTCTTTG...GTTAGTCAGCTGGTACGAAAAGTTGGGGAAGGA 1159

 299 PhePheValAspGlu..................................... 303
     :::      :::|||
1160 TATGTATTCGAAGAAAAGGGCATCTCTCGTTATGTCTTTGCGAAAGATTT 1209

 304 ....SerGluProLeuLeuArgCysAspSerThrSerSerGlySerSerA 319
         ||||||   :::   :::   :::|||      |||   :::::::
1210 ACCATCTGAAACTGTTAAAAATCTTGAAAGCAAGTTATCAAAACAAGAGA 1259

 319 laLeuSer.......................................... 321
     :::::|||
1260 GTGTTTCACACACTTTAACTGCTAAAAAAGAAAATGTTGCTCCTCGTGAC 1309

 321 ..................................................... 321

1310 CAAGAATTTTATGATAAAGCATATAATCTGTTAACTGAGGCTCATAAAGC 1359

 322 ..................ArgAsnGlySerPheIleThrLysGluLysL 332
                        ||||||::::::|||         :::
1360 CTTGTTTGNAAATAAGGGTCGTAATTCTGATTTC.........CAAGCCT 1400

 332 ysAspThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnPro 348
       |||   :::|||:::::::::::::   :::       :::   :::
1401 TAGACAAATTATTAGAACGCTTGAATGATGAATCGACTAATAAAGAAAAA 1450

 349 IlePheAspAspMetLeuHisPheLeu.............AsnProGluGl 361
     :::   ||||||:::|||   ||||||            :::|||||::
1451 TTGGTAGATGATTTATTGGCATTCCTAGCACCAATTACCCATCCAGAGCG 1500

 361 uLeu...ArgValIleGluGluIleProGlnAlaGluAspLysLeu 375
     :|||   :::       ::::::|||      :::||||||::::::
1501 ACTTGGCAAACCAAATTCTCAAATTGAGTATACTGAAGACGAAGTT 1546
```

```
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-985-916-15

seq_documentation_block:
; Sequence 15, Application US/08985916
; Patent No. 6221636
;  GENERAL INFORMATION:
;    APPLICANT:  ATSUSHI HAYAKAWA, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSH
;    TITLE OF INVENTION:  METHOD FOR PRODUCING L-LYSINE
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
;      CITY:  ARLINGTON
;      COUNTRY:  VA
;      ZIP:  22152
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,916
;      FILING DATE:  05-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 8-325658
;      FILING DATE:  05-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  NORMAN F. OBLON
;      REGISTRATION NUMBER:  24,618
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3533 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Brevibacterium lactofermentum
;      STRAIN:  ATCC 13869
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  321..3077
US-08-985-916-15


alignment_scores:
            Quality:  101.00              Length:    370
              Ratio:  0.580                Gaps:     25
 Percent Similarity:  47.027   Percent Identity: 24.054

alignment_block:
 US-09-733-956-2 x US-08-985-916-15   ..

 Align seg 1/1  to: US-08-985-916-15  from: 1  to: 3533

      54 ProHisHisArgHisIleLeuLysLeuLeuProSerMetGluAlaThrGl 70
         |||   ||||||   :::   :::   :::   |||   ||||||   ||
    1607 CCAAGTCACCGCAAACTACCGCGAGCTGTCTGAAGCAGAGAAGCTTGAGG 1656

      70 y...GlyGluLysSerSerThrProIleLysGlyPro..............  81
         |   :::|||   ::::::   |||   ::::::
    1657 TGCTGCTGAAGGAACTGCGCAGCCCTCGTCCGCTGATCCCGCACGGTTCA 1706

      82 ........LysArgGlyHis...ProArgGlnAsnLeuHisLysHis.Ph  94
                :::|||||||||   ||||||            :::||| ||
    1707 GATGAATACAGCGAGGTCACCGACCGCGAGCT.........CGGCATCTT 1747

      94 eAspIleAsnGluHisLeuProTrpMetIleValLeuPheLeuLeuLeuV 111
         |
    1748 C.................................................. 1748

     111 alLeuValValIleValValCysSerIleArgLysSerSerArgThrLeu 127
                                  |||   :::|||:::::::::
    1749 ............................CGCACCGCGTCGGAGGCTGTT 1769

     128 LysLys...GlyProArgGlnAspProSerAlaIleValGluLysAlaGl 143
         ||||||   ||||||||||::   |||   :::|||:::::   |||
    1770 AAGAAATTCGGGCCACGGATGGTGCCTCACTGCATCATCTCCATGGCA.. 1817

     143 yLeuLysLysSerMetThrProThrGlnAsn..............ArgG 155
                :::|||:::|||   :::   :::               :::|
    1818 .......TCATCGGTCACCGATGTGCTCGAGCCGATGGTATTGCTCAAGG 1860

     155 luLysTrpIleTyrTyrCysAsnGlyHisGly............IleAsp 167
         ||      :::      :::||||||   :::              :::|||
    1861 AATTCGGCCTCATTGCAGCCAACGGCGACAACCCACGCGGCACCGTCGAT 1910

     168 IleLeuLysLeuVal.....................................AlaAlaGl 175
         ::::::   |||                                    |||:::
    1911 GTCATCCCACTGTTCGAAACCATCGAAGATCTCCAGGCCGGCGCCGGAAT 1960

     175 nValGlySerGlnTrpLys...AspIleTyrGln......PheLeuCysA 189
           :::   :::   ||||||   |||:::|||:::      :::|||   :
    1961 CCTCGACGAACTGTGGAAAATTGATCTTTACCGCAACTACCTCCTGCAGC 2010

     189 snAlaSerGluArgGluValAlaAlaPheSerAsnGlyTyrThrAlaAsp 205
         ::   :::   :::|||||||              ||||||:::   :::
    2011 GCGACAACGTCCAGGAAGTCATGCTC.........GGTTACTCCGATTCC 2051

     206 HisGluArgAlaTyrAlaAlaLeuGlnHisTrpThrIleArgGlyProGl 222
         ::::::   :::           :::|||::::::          ||
    2052 AACAAGGATGGCGGATATTTCTCCGCAAACTGGGCGCTTTACGACGCGGA 2101

     222 uAlaSerLeuAlaGlnLeuIleSerAla.............LeuArgGln. 234
         |   :::|||::::::|||      :::              ||||||
    2102 ACTGCAGCTCGTCGAACTATGCCGATCAGCCGGGGTCAAGCTTCGCCTGT 2151

     235 ..HisArgArgAsnAspValValGluLys.IleArgGlyLeuMetGluAs 250
           |||   |||:::   :::|||  ::: :::      |||
    2152 TCCACGGCCGTGGTGGCACCGTCGGCCGCGGTGGCGGACCTT......CC 2195

     250 pThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerP 267
          ||||||::::::                  ||||||   ||||||:::
    2196 TACGACGCGATTC.................TTGCCC...AGCCCAGGGG 2224

     267 roLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSer 283
         ||||||      |||      ||||||   ::::::   ::::::|||
    2225 GGCTGTCCAAGGTTCCGTGCGCATCACCG...AGCAGGGCGAGATCATCT 2271

     284 AlaLeuLeuThrValGluProSerPro..................... 292
           |||   |||:::   |||:::|||
    2272 CCGCTAAGTACGGCAACCCCGAAACCGCGCGCCGAAACCTCGAAGCTCTG 2321

     293 .......GlnAspLysAsnLysGlyPhePheValAspGluSerGluProL 307
              |||      :::      |||   ::::::   ::::::|||:
    2322 GTCTCAGCAACGCTTGAGGCATCGCTTCTCGACGTCTCCGAACTCACCGA 2371

     307 euLeuArgCysAspSerThrSer.........SerGlySerSerAlaLeu 320
         ::   ::::::   :::||||||          |||   |||||||||
    2372 TCACCAACGCGCGTACGACATCATGAGTGAGATCTCTGAGCTCAGCTTGA 2421

     321 SerArgAsnGlySerPheIleThrLysGluLysLysAspThrValLeu.. 336
              :::      :::   |||:::   |||   :::::::::
    2422 AGAAGTACGCCTCCTTGGTGCACGAGGATCAAGGCTTCATCGATTACTTC 2471

     337 .............ArgGlnValArgLeuAspProCysAspLeuGlnProI 349
                      |||      ||||||||||||          :::|||
    2472 ACCCAGTCCACGCCGCTGCAGGAGATTGGATCCCTCAACATCGGATCCA. 2520

     349 lePheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIle 365
                    :::   |||      :::|||   :::   ::::::
    2521 ........GGCCTTCCTCACGCAAGCAGACCTCCTCGGTGGAAGATTTGC 2562

     366 GluGlu 367
         |||:::
    2563 GAGCAA 2568

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-324-867-1

seq_documentation_block:
; Sequence 1, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:
;  APPLICANT: Lillicrap, David
;  APPLICANT: Cameron, Cherie
;  APPLICANT: No. 6251632ley, Colleen
;  APPLICANT: Horrocks, L. Suzanne Hoyle
;  APPLICANT: Hough, Christine
;  TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
;  FILE REFERENCE: 1669.0010002/JAG/BJD
;  CURRENT APPLICATION NUMBER: US/09/324,867A
;  CURRENT FILING DATE: 1999-06-03
;  EARLIER APPLICATION NUMBER: 09/035,141
;  EARLIER FILING DATE: 1998-03-059
;  EARLIER APPLICATION NUMBER: 60/039,953
;  EARLIER FILING DATE: 1997-03-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 7032
;   TYPE: DNA
;   ORGANISM: Canis familiaris
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..7029
US-09-324-867-1


alignment_scores:
            Quality:   99.00              Length:    458
```

```
            Ratio:  0.503                Gaps:     19
 Percent Similarity:  43.013   Percent Identity: 19.432

alignment_block:
 US-09-733-956-2 x US-09-324-867-1

 Align seg 1/1  to: US-09-324-867-1  from: 1  to: 7032

      10 AlaSerValArgProLysValLeuSerSerIleGlnGluGlyThrValPr 26
         ||||||:::||||||:::::      |||   ::::::      :::||
    2536 GCAAGTCTCAGACCAGAGCTCCGTCACAGTGAGGACAGAGAATTTACTCC 2585

      26 oAspAsnThrSerSerAlaArgGlyLysGluAspValAsnLysThrLeuP 43
         |:::                                             |
    2586 TGAG.............................................C 2590

      43 roAsnLeuGlnVal...ValAsnHisGlnGlnGlyProHisHisArgHis 58
         ||:::||||||:::   :::|||::::::   |||   :::
    2591 CAGAACTGCAGTTAAGATTAAATGAGAATTTGGGGACAAATACAACAGTA 2640

      59 IleLeuLys.....................................LeuLe 63
            ||||||                                    |||::
    2641 GAGTTGAAGAAACTTGATTTAAAAATTTCTAGTTCATCAGACAGTCTAAT 2690

      63 uProSerMetGluAlaThrGlyGlyGluLysSerSerThrProIleLysG 80
         :   |||            :::   :::   :::::::::      ||
    2691 GACTTCACCAACAATTCCATCAGATAAGTTGGCAGCAGCTACTGAAAAGA 2740

      80 lyProLysArgGlyHisProArgGlnAsnLeuHisLysHisPheAspIle 96
            :::   |||   |||:::::::::::::|||           :::
    2741 CAGGTTCCTTAGGACCCCCAAATATGTCAGTTCAC.....TTT 2778

      97 AsnGluHisLeuProTrpMetIleVal..................... 105
         |||:::||||||      ::::::
    2779 AACAGTCATTTAGGTACCATTGTATTTGGCAATAATTCATCCCACCTTAT 2828

     106 .............................................LeuP 107
                                                       :
    2829 TCAGTCTGGTGTACCTTTGGAATTGAGTGAAGAAGATAATGATTCCAAGT 2878

     107 heLeuLeuLeuValLeuValValIleValValCysSerIleArg..... 121
         ::|||         |||:::   |||         |||:::|||
    2879 TGTTAGAAGCACCTTTAATGAATATTCAAGAAAGTTCACTGAGAGAAAAT 2928

     122 ............LysSerSerArgThrLeuLysLysGlyProArgGlnAs 134
                     :::|||:::|||   :::|||:::       :::
    2929 GTATTATCAATGGAGAGTAATAGGTTATTTAAAGAAGAAAGAATTCGTGG 2978

     134 pProSerAlaIleValGluLys.....................AlaG 143
          |||:::::::::::::::                        ::::
    2979 ACCTGCTTCATTAATCAAAGATAATGCTTTATTCAAAGTTAATATCTCTT 3028

     143 lyLeuLysLysSerMetThrPro.........ThrGlnAsnArgGluLys 156
         :::::|||   :::   :::|||         |||   ||||||:::
    3029 CGGTAAAGACAAACAGGGCACCAGTTAACTTAACAACTAATAGAAAGACT 3078

     157 TrpIleTyrTyrCysAsnGlyHisGlyIleAspIleLeuLysLeuValAl 173
                              :::   |||      |||:::
    3079 .....................CGTGTTGCTATCCCAACATTATTAAT 3104

     173 aAlaGlnValGlySerGlnTrpLysAspIleTyrGlnPheLeuCysAsnA 190
            :::      |||   |||:::||||||   ::::::      |||:
    3105 TGAGAACAGTACCTCAGTCTGGCAAGATATT...ATGTTAGAAAGGAATA 3151

     190 laSerGluArgGluValAlaAlaPheSerAsnGlyTyrThr...AlaAsp 205
         :::::   :::||||||:::::::::   ::::::   |||      |||
    3152 CTGAGTTTAAAGAAGTAACTTCTTTGATTCATAATGAAACGTTTATGGAC 3201

     206 HisGluArgAlaTyrAlaAlaLeuGlnHisTrpThrIleArgGlyProGl 222
         ::::::   :::      :::|||:::|||   :::             ::
    3202 AGAAATACTACAGCTCTGGGGCTAAATCATGTGTCA.....AATAA 3242

     222 uAlaSerLeuAlaGlnLeuIleSerAlaLeuArgGlnHisArgArgAsnA 239

         :::::::|||::::::   ::::::      :::|||   :::
    3243 AACTACTTTATCAAAAAATGTGGAAATGGCCCACCAAAAAAAA..... 3285

     239 spValValGluLysIleArgGlyLeuMetGluAspThrThrGlnLeuGlu 255
                                      ||||||   :::   |||:::
    3286 ............................GAAGACCCTGTGCCACTACGT 3306

     256 ThrAspLysLeuAlaLeuProMetSerProSerProLeuSerProSerPr 272
         :::::::::      |||      |||      |||:::   |||:::
    3307 GCAGAAAATCCAGATCTATCATCCTCCAAGATACCGTTCTTGCCAGATTG 3356

     272 oIleProSerProAsnAlaLysLeuGluAsnSerAlaLeuLeuThrValG 289
          |||   :::   :::          :::||||||              :
    3357 GATAAAGACCCATGGC.........AAGAACTCCCTAAGCTCTGAGCAAA 3397

     289 luProSerProGlnAspLysAsnLysGlyPhePheValAspGluSerGlu 305
         ::||||||||||:::
    3398 GGCCCAGTCCAAAA..................................... 3411

     306 ProLeuLeuArgCysAspSerThrSerSerGlySerSerAlaLeuSerAr 322
                        :::   ||||||   |||||||:::         ::
    3412 ...............CAATTAACATCTTTAGGATCAGAAAAATCTGTGAA 3446

     322 gAsnGlySerPheIleThrLysGluLys..................... 331
         ::::   :::|||::::::::::||||||
    3447 AGATCAGAACTTTTTGTCAGAGGAGAAGGTGGTAGTAGGAGAGGATGAAT 3496

     332 .....LysAspThrValLeuArgGlnValArgLeuAspProCysAspLeu 346
               |||||||||                              :::|||
    3497 TTACGAAGGACACA.............................GAACTC 3516

     347 GlnProIlePhe......AspAspMetLeuHisPheLeuAsn........ 358
         |||   ||||||      :::      :::   |||   |||
    3517 CAAGAGATTTTTCCAAACAACAAGAGCATATTTTTTGCTAACTTGGCTAA 3566

     359 .....................................ProGluGluL 362
                                              |||||||||:
    3567 TGTCCAAGAAAATGATACATACAATCAAGAAAAAAAATCTCCGGAAGAGA 3616

     362 euArgValIleGluGluIleProGlnAlaGluAspLysLeuAspArgLeu 378
         :::::      |||:::::   |||   :::      |||   :::
    3617 TAGAAAGAAAGGAAAAATTAACCCAGGAGAATGTGGCTTTGCCTCAGGCA 3666

     379 PheGluIleIleGlyValLysSer 386
               :::||||||:::|||:::
    3667 CATACTATGATTGGCACTAAGAAC 3690

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-989-925-2

seq_documentation_block:
; Sequence 2, Application US/08989925
; Patent No. 5989820
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  HUMAN ADIPOPHILIN-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/989,925
```

```
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0440 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1947 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRAITUT13
;      CLONE:  1620223
US-08-989-925-2


alignment_scores:
            Quality:   96.50              Length:    422
              Ratio:  0.524                Gaps:     18
 Percent Similarity:  43.602   Percent Identity: 20.142

alignment_block:
 US-09-733-956-2 x US-08-989-925-2   ..

 Align seg 1/1  to: US-08-989-925-2  from: 1  to: 1947

       4 ThrGluSerAsnSerSerAlaSerValArgProLysValLeuSerSerIl 20
         |||   :::   :::::::::|||           :::||||||:::::
     242 ACCCTCACGGCGGCTGCTGTCAGCGGGGCTCAGCCGATCCTCTCCAAGCT 291

      20 eGlnGluGlyThrValProAspAsnThrSerSerAlaArgGlyLysGluA 37
         ::::       :::  :::      :::   ||||||   :::
     292 GGAGCCCCAGATTGCATCAGCCAGCGAATACGCCCACAGGGGGCTGGACA 341

      37 spValAsnLysThrLeuProAsnLeuGlnValValAsnHisGlnGlnGly 53
         :::::::::::||||||   ||||||         |||   :::|||
     342 AGTTGGAGGAGAACCTCCCCATCCTGCAGCA....GCCCACGGAGAAGGT 387

      54 ProHisHisArgHisIleLeuLysLeuLeuProSerMetGluAlaThrGl 70
         |||                                          ||
     388 CCT.............................................GG 392

      70 yGlyGluLysSerSerThr.........ProIleLysGlyProLysArgG 84
         ||||:::::::::::::              :::  ||||||:::   |
     393 CGGACACCAAGGAGCTTGTGTCGTCTAAGGTGTCGGGGGCCCAAGAGATG 442

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||      ||||||:::         |||
     443 GTGTCTAGCGCCAAGGACACGGTGGCCAC..................... 471

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117

     471 ................................................. 471

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
                                 :::     |||     :::|
     472 .....................CCAATTGTCGGAGGCGGTGGACGCG 496

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         |||||   :::   :::|||:::   |||     |||      :::
     497 ACCCGCGGTGCTGTGCAGAGCGGCGTGGACAAGACAAAGTCCGTAGT... 543

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
               :::|||::::::   |||            |||||||||:::

     544 ......GACCGGCGGCGTCCAATC...........GGTCATGGGCTCCC 575

     167 p.IleLeuLysLeuValAlaAlaGlnVal.................Gly 177
            :::   ::::::|||   :::   |||                 :::
     576 GCTTGGGCCAGATGGTGTTGAGTGGGGTCGACACGGTGCTGGGGAAGTCG 625

     178 SerGlnTrpLysAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGl 194
         :::::|||   |||   :::      |||   :::|||:::      ::
     626 GAGGAGTGGGCGGACAACCACCTGCCCCTTACGGATGCCGAACTGGCCCG 675

     194 uValAlaAlaPheSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrA 211
         ::::|||:::      :::|||:::   :::         :::
     676 CATCGCCACATCCCTGGATGGCTTCGACGTTGCGTCCGTGCAGCAGCAGC 725

     211 laAlaLeuGlnHisTrpThrIleArgGlyProGluAlaSerLeuAlaGln 227
                |||   :::   :::|||               |||::::::
     726 GGCAGGAACAGAGCTACTTCGTACGT..............CTGGGCTCC 760

     228 LeuIleSerAlaLeuArgGlnHisArgArgAsnAspValValGluLysIl 244
         |||   :::   ||||||||||||      :::      :::   |||::
     761 CTGTCGGAGAGGCTGCGGCAGCACGCCTATGAGCACTCGCTGGGCAAGCT 810

     244 eArg.............................................. 245
         :|||
     811 TCGAGCCACCAAGCAGAGGGCACAGGAGGCTCTGCTGCAGCTGTCGCAGG 860

     246 .....GlyLeuMetGluAspThrThrGlnLeuGluThrAspLysLeuAla 260
               :::||||||||||   :::   |||         :::||||||:::
     861 CCCTAAGCCTGATGGAAACTGTCAAGCAAGGCGTTGATCAGAAGCTGGTG 910

     261 ..........................LeuProMetSerProSerPr 267
                                     |||      :::   :::
     911 GAAGGCCAGGAGAAGCTGCACCAGATGTGGCTCAGCTGGAACCAGAAGCA 960

     267 oLeu...SerProSerProIleProSerProAsnAlaLysLeuGluAsnS 283
          |||   :::|||:::      |||             :::::::|||:::
     961 GCTCCAGGGCCCCGAGAAGGAGCCGCCCAAGCCAGAGCAGGTCGAGTCCC 1010

     283 erAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
           ||||||                                       :::
    1011 GGGCGCTCACC.....................................ATG 1024

     300 PheValAspGluSerGluProLeu...LeuArgCysAspSerThrSerSe 315
         |||   |||   ::::::   |||          |||   |||   :::||
    1025 TTCCGGGACATTGCCCAGCAACTGCAGGCCACCTGTACCTCCCTGGGGTC 1074

     315 rGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysL 332
         |:::   :::::::|||      |||            :::|||::::::
    1075 CAGCATTCAGGGCCTCCCCACCAAT............GTGAAGGACCAGG 1112

     332 ysAspThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnPro 348
           :::   :::   |||||||||:::            |||||||||
    1113 TGCAGCAGGCCCGCCGCCAGGTGGAG..........GACCTCCAGGCC 1150

     349 IlePheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIl 365
         |||::::::   :::|||   ::::::      :::      :::::
    1151 ACGTTTTCCAGC...ATCCACTCCTTCCAGGACCTGTCCAGCAGCATTCT 1197

     365 e...........GluGluIleProGlnAlaGluAspLysLeuAspArgL 378
         :           |||::::::   :::|||::::::   ||||||::::
    1198 GGCCCAGAGCCGTGAGCGTGTCGCCAGCGCCCGCGAGGCCCTGGACCACA 1247

     378 euPheGluIleIle 382
         ::   |||   :::
    1248 TGGTGGAATATGTG 1261
```

```
OM of: US-09-733-956-2 to: N_Geneseq_1101:*   out_format : pfs

Date: Jan 4, 2002  9:11 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09733956/runat_04012002_084026_15566/app_query.fasta_1.467
 -DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09733956_@CGN1_1_305 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-733-956-2
Query length: 405
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 163.610000

score_list:
Sequence             Strd Orig    ZScore   EScore Len  ! Documentation
/SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV57441 + 2091.00 3492.71  1.8e-186  2186  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA28728 + 2091.00 3492.29  1.9e-186  2271  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ57184 + 2091.00 3490.53  2.4e-186  2666  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ92406 + 2091.00 3489.69  2.7e-186  2877  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA47396 + 2091.00 3488.39  3.1e-186  3236  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA47395 + 2091.00 3488.25  3.2e-186  3277  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25259 + 2091.00 3487.71  3.4e-186  3440  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV99927 + 2091.00 3487.61  3.5e-186  3474  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA99658 + 2091.00 3487.61  3.5e-186  3474  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ33945 + 2091.00 3487.42  3.6e-186  3534  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC78474 + 2091.00 3487.42  3.6e-186  3534  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH14374 + 2087.00 3483.91  5.6e-186  2636  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25320 + 1914.00 3190.03  1.3e-169  3331  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25264 + 1567.00 2611.97  2.1e-137  1893  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25322 + 1527.00 2540.96  1.9e-133  2638  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA28727 + 1511.00 2515.09  5.1e-132  2393  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX89504 + 1352.00 2252.69  2.1e-117  1476  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25261 + 1322.00 2195.57  3.2e-114  2686  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25321 + 1269.00 2106.64  2.9e-109  2612  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25260 + 1269.00 2106.49  3.0e-109  2646  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ92407 +  992.50 1647.36  1.1e-83   1474  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI12543 +  804.00 1342.63  1.0e-66   467  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI33893 +  804.00 1342.63  1.0e-66   467  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI02448 +  804.00 1342.63  1.0e-66   467  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV99929 +  579.00  966.10  9.8e-46   378  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA99660 +  579.00  966.10  9.8e-46   378  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV99928 -  564.00  941.23  2.4e-44   365  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA99659 -  564.00  941.23  2.4e-44   365  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV87597 +  450.00  751.96  8.3e-34   286  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI21181 +  439.00  734.53  7.7e-33   259  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI22711 +  439.00  734.53  7.7e-33   259  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI46449 +  439.00  734.53  7.7e-33   259  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI48007 +  439.00  734.53  7.7e-33   259  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI06896 +  439.00  734.53  7.7e-33   259  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI08374 +  439.00  734.53  7.7e-33   259  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV99931 +  398.50  664.15  6.5e-29   316  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA99662 +  398.50  664.15  6.5e-29   316  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI22923 +  397.00  665.22  5.6e-29   228  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI48220 +  397.00  665.22  5.6e-29   228  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI08587 +  397.00  665.22  5.6e-29   228  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI11976 +  373.00  616.95  2.7e-26   465
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI33308 +  373.00  616.95  2.7e-26   465
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI01909 +  373.00  616.95  2.7e-26   465
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI13702 +  349.00  579.24  3.5e-24   364
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI35059 +  349.00  579.24  3.5e-24   364

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV57441

seq_documentation_block:
ID   AAV57441 standard; cDNA; 2186 BP.
XX
AC   AAV57441;
XX
DT   21-DEC-1998  (first entry)
XX
DE   Human tumour necrosis factor related receptor TR7 cDNA.
XX
KW   Tumour necrosis factor related receptor; TR7; human; inflammation;
KW   arthritis; septicaemia; autoimmune disease; psoriasis;
KW   inflammatory bowel disease; infection; graft-versus-host disease;
KW   transplant rejection; stroke; acute respiratory disease syndrome;
KW   ischaemia; restenosis; brain injury; AIDS; bone disease; cancer;
KW   atherosclerosis; Alzheimer's disease; diagnosis; therapy; vaccine;
KW   ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             7..1974
FT                   /*tag=  a
FT                   /transl_except= (pos:76..78, aa:Arg)
XX
PN   EP869179-A1.
XX
PD   07-OCT-1998.
XX
PF   01-APR-1998;   98EP-0302528.
XX
PR   28-OCT-1997;   97US-0959382.
PR   02-APR-1997;   97US-0041796.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Deen KC,  Hurle MR,  Tan KB,  Young P;
XX
DR   WPI; 1998-508493/44.
DR   P-PSDB; AAW75792.
XX
PT   New tumour necrosis factor receptor TR7 polypeptides and
PT   polynucleotides - useful as diagnostic reagents and for treating
PT   Alzheimer's disease, AIDS and cancer
XX
PS   Claim 4; Page 18-19; 25pp; English.
XX
CC   This cDNA clone codes for a novel member of the human tumour
CC   necrosis factor (TNF) receptor superfamily, termed TR7 (see AAW75792).
CC   3 ESTs with sequence similarity to the human TNF receptor were
CC   discovered in a commercial EST database.  Analysis of 2 of these
CC   revealed that each was a partial sequence of the complete coding
CC   region, overlapping, with 100% identity.  Together, the 2 clone
CC   sequences encompassed the complete predicted coding sequence of TR7
CC   of 1968 bp.  TR7 polynucleotides are useful for diagnosing
CC   diseases or susceptibility to diseases by identifying mutations in
CC   the TR7 gene using probes containing the TR7 nucleotide sequence.
CC   TR7 polypeptide can be used to screen for agonists and antagonists
CC   which bind the receptor.  Gene therapy may be used to effect
CC   endogenous TR7 polypeptide production.  TR7 polypeptides and
CC   polynucleotides can be administered directly as vaccines for
CC   prevention of diseases.  Diseases diagnosed, treated or prevented
CC   by the above methods include: chronic and acute inflammation,
CC   arthritis, septicemia, autoimmune diseases (e.g. inflammatory bowel
CC   disease, psoriasis), transplant rejection, graft vs. host disease,
CC   infection, stroke, ischaemia, acute respiratory disease syndrome,
```

```
CC   restenosis, brain injury, AIDS, bone diseases, cancer (e.g.
CC   lymphoproliferative disorders), atherosclerosis, and Alzheimer's
CC   disease.  The TR7 polynucleotide is also useful for mapping the
CC   gene to a chromosome, allowing gene inheritance to be studied
CC   through linkage analysis.
XX
SQ   Sequence 2186 BP; 514 A; 613 C; 565 G; 494 T; 0 other;


alignment_scores:
           Quality: 2091.00              Length:    405
             Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAV57441   ..

 Align seg 1/1  to: AAV57441  from: 1  to: 2186

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     757 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 806

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     807 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 856

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     857 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 906

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     907 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 956

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     957 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1006

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1007 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1056

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1057 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1106

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1107 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1156

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1157 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1206

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1207 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1256

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1257 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1306

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1307 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1356

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1357 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1406

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||

    1407 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1456

      234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1457 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1506

      251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1507 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1556

      267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1557 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1606

      284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1607 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1656

      301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1657 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1706

      317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1707 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1756

      334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1757 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 1806

      351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1807 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 1856

      367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1857 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 1906

      384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1907 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 1956

      401 LeuProAspLeuLeu 405
          ||||||||||||||||
    1957 CTTCCTGACCTGCTG 1971

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA28728

seq_documentation_block:
ID   AAA28728 standard; DNA; 2271 BP.
XX
AC   AAA28728;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Human osteoprotegrin-like 2 (OPGx2) DNA.
XX
KW   OPGx1; osteoprotegrin-like; DR6 TNF-related death receptor; agonist;
KW   inhibitor; bone resorption; vascular calcification; apoptosis;
KW   osteopathic; chromosome 6p11.1; D6S452; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             102..2177
FT                   /*tag=  a
FT                   /product= OPGx2
XX
PN   WO200024771-A2.
XX
PD   04-MAY-2000.
XX
PF   22-OCT-1999;   99WO-US24913.
```

```
XX
PR   23-OCT-1998;   98US-0105481.
PR   01-OCT-1999;   99US-0156993.
PR   21-OCT-1999;   99US-0422680.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Yang M,  Lichenstein H,  McDonald WF;
XX
DR   WPI; 2000-350692/30.
DR   P-PSDB; AAY92846.
XX
PT   Nucleic acids encoding osteoprotegerin-like (OPGx) polypeptides, useful
PT   for treating disorders associated with bone metabolism, such as
PT   osteoporosis and osteopetrosis
XX
PS   Claim 1; Fig 2; 81pp; English.
XX
CC   Osteoprotegrin-like polypeptides (OPGx) are expressed in bone, lymph
CC   node, germinal B cells and kidney. There appear to be at least two splice
CC   variants of this gene, a transmembrane form and an extracellular domain
CC   form. Human OPGx1, 2 and 4 contain a 36 amino acid extension at their
CC   N-terminal relative to the DR6 tumour necrosis factor (TNF)-related death
CC   receptor. The OPGx gene has been mapped to chromosome 6p11.1, essentially
CC   on top of D6S452. The OPGx polypeptides, agonists and antibodies are
CC   useful in methods to inhibit osteoclast-mediated bone resorption or
CC   vascular calcification and to modulate cell death (apoptosis). This is
CC   useful for treating disorders associated with bone metabolism, such as
CC   osteoporosis, osteopetrosis, or a condition characterized by loss of
CC   bone, breakdown of tissue, or excessive readsorption of bone tissue.
XX
SQ   Sequence 2271 BP; 518 A; 668 C; 611 G; 467 T; 7 other;


alignment_scores:
            Quality: 2091.00              Length:    405
              Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAA28728   ..

 Align seg 1/1  to: AAA28728  from: 1  to: 2271

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     960 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1009

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1010 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1059

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1060 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1109

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1110 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1159

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1160 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1209

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1210 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1259

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1260 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1309

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1310 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1359

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1360 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1409

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1410 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1459

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1460 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1509

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1510 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1559

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1560 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1609

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1610 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1659

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1660 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1709

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1710 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1759

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1760 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1809

     284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1810 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1859

     301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1860 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1909

     317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1910 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1959

     334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1960 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2009

     351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2010 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2059

     367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2060 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2109

     384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2110 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 2159

     401 LeuProAspLeuLeu 405
         |||||||||||||||
    2160 CTTCCTGACCTGCTG 2174

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ57184
```

```
seq_documentation_block:
ID   AAZ57184 standard; cDNA; 2666 BP.
XX
AC   AAZ57184;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Human DETH protein encoding cDNA SEQ ID NO:1.
XX
KW   Human; DETH; apoptosis; TNFR; neurodegenerative diseases; cancer;
KW   death domain expressing tumour necrosis factor receptor family homologue;
KW   autoimmune disease; diagnosis; ss.
XX
OS   Homo sapiens.
XX
PN   WO9966039-A1.
XX
PD   23-DEC-1999.
XX
PF   08-JUN-1999;   99WO-GB01793.
XX
PR   12-JUN-1998;   98GB-0012607.
XX
PA   (ZENE ) ZENECA LTD.
XX
PI   Lu JJ,  Gomes BC,  Fieles WE;
XX
DR   WPI; 2000-097744/08.
DR   P-PSDB; AAY67947.
XX
PT   New protein having 80% identity to death domain expressing tumor
PT   necrosis factor receptor family homologue -
XX
PS   Claim 7; Page 33-34; 40pp; English.
XX
CC   The present sequence encodes a human death domain expressing tumour
CC   necrosis factor receptor family homologue, designated DETH. DETH has a
CC   role in apoptosis and appears to be member of the TNFR1/FAS/NGFR family
CC   of receptors. The DETH protein can be used for inducing apoptosis by
CC   expressing DETH in a cell. The protein is useful for identifying
CC   inhibitors and agonists of apoptosis that treat neurodegenerative
CC   diseases (inhibition); cancer and autoimmune diseases (induction).
CC   DETH-specific antibodies are useful for diagnosis of conditions and
CC   diseases associated with DETH expression. The protein avoids the use
CC   of harmful chemotherapy, where the loss of the p53 tumour repressor
CC   gene can lead to drug resistant tumour cells following treatment.
XX
SQ   Sequence 2666 BP; 691 A; 609 C; 591 G; 775 T; 0 other;


alignment_scores:
            Quality: 2091.00              Length:    405
              Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAZ57184   ..

 Align seg 1/1  to: AAZ57184  from: 1  to: 2666

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     190 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 239

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     240 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 289

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     290 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 339

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     340 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 389

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     390 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 439

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     440 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 489

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     490 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 539

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     540 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 589

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     590 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 639

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     640 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 689

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     690 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 739

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     740 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 789

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     790 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 839

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     840 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 889

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     890 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 939

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     940 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 989

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     990 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1039

     284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1040 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1089

     301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1090 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1139

     317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1140 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1189

     334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1190 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 1239

     351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
    1240 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 1289

     367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1290 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 1339

     384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1340 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 1389

     401 LeuProAspLeuLeu 405
         ||||||||||||||||
    1390 CTTCCTGACCTGCTG 1404

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ92406

seq_documentation_block:
ID   AAZ92406 standard; cDNA; 2877 BP.
XX
AC   AAZ92406;
XX
DT   05-JUN-2000  (first entry)
XX
DE   cDNA encoding human TNF receptor-like protein HSLJD37R, SEQ ID NO:9.
XX
KW   TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
KW   Rank-like protein; RANKL; immune disorder; inflammation; allergy;
KW   immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
KW   dermatological; antithyroid; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             410..2377
FT                   /*tag= a
FT                   /product= "Human HSLJD37R"
XX
PN   WO200001817-A2.
XX
PD   13-JAN-2000.
XX
PF   06-JUL-1999;   99WO-US12366.
XX
PR   06-JUL-1998;   98US-0110938.
PR   13-JUL-1998;   98US-0114466.
PR   23-JUL-1998;   98US-0093897.
PR   12-AUG-1998;   98US-0132968.
PR   18-AUG-1998;   98US-0136214.
PR   11-SEP-1998;   98US-0099999.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bates EEM,  Lebecque SJE,  Murphy EE,  Mattson JD,  Gorman DM;
PI   Hedrick JA,  Wang L,  Zlotnik A,  Murgolo NJ,  Greene JR,  Johnston JA;
PI   Bazan JF,  Mahony D,  Lees EM;
XX
DR   WPI; 2000-171015/15.
DR   P-PSDB; AAY77460.
XX
PT   New isolated mammalian genes, used to develop products for treating
PT   e.g. immune, inflammatory or allergic abnormalities, cancers or
PT   degenerative conditions  -
XX
PS   Claim 25; Page 160-163; 218pp; English.
XX
CC   The invention relates to a number of primate and/or rodent proteins, and
CC   the genes which encode them. The invention encompasses human dendritic
CC   cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC   factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
CC   human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub
CC   12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC   properties of ligands for proteins comprising a leucine-rich motif
CC   (LRR); human cyclin E2; cDNAs encoding these proteins; and antibodies
CC   against these proteins. The proteins can be used for modulating the
CC   physiology or development of a cell. They can be used to mediate uptake
CC   of substrates (e.g., prostaglandin-like molecules), to modulate or
CC   mediate cellular interactions (e.g., induce or prevent trafficking,
CC   proliferation, or differentiation of cells), or are intracellular
CC   proteins which are important in various cellular processes such as the
CC   deubiquitination of proteins or cell cycle regulation. The products can
CC   be used for treating medical conditions such as immune, inflammatory or
CC   allergic disorders, or abnormal cellular proliferation, for example,
CC   cancers or degenerative conditions. They can be used to modulate immune
CC   responses in disease states e.g., autoimmune disorders, including
CC   rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's
CC   autoimmune thyroiditis, as well as acute and chronic inflammatory
CC   responses in which T cell activation, expansion, and/or immunological T
CC   cell memory play an important role. Sequences AAZ92404-Z92411 represent
CC   cDNAs encoding TNF receptor family-related proteins. AAZ92404 encodes the
CC   human protein HDTEA84, AAZ92405-Z92407 encode human HSLJD37R proteins,
CC   AAZ92408 encodes murine Rank-like protein RANKL, and AAZ92409-Z92411
CC   encode human RANKL proteins.
XX
SQ   Sequence 2877 BP; 636 A; 817 C; 751 G; 673 T; 0 other;


alignment_scores:
                Quality: 2091.00              Length:    405
                  Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAZ92406   ..

 Align seg 1/1  to: AAZ92406  from: 1  to: 2877

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1160 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1209

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1210 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1259

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1260 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1309

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1310 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1359

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1360 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1409

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1410 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1459

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1460 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1509

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1510 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1559

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1560 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1609

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1610 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1659
```

```
  167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1660 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1709

  184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1710 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1759

  201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1760 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1809

  217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1810 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1859

  234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1860 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1909

  251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1910 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1959

  267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1960 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 2009

  284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2010 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 2059

  301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2060 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 2109

  317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2110 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 2159

  334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2160 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2209

  351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2210 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2259

  367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2260 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2309

  384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2310 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 2359

  401 LeuProAspLeuLeu 405
      ||||||||||||||||
 2360 CTTCCTGACCTGCTG 2374

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA47396

seq_documentation_block:
ID   AAA47396 standard; cDNA; 3236 BP.
XX
AC   AAA47396;
XX
DT   20-OCT-2000  (first entry)
XX
DE   Tumour necrosis factor receptor homologue TRH1 (Clone 2098183).
XX
KW   Tumour necrosis factor receptor homologue; TRH1; TNF; arthritis;
KW   transplant rejection; activation; proliferation; differentiation;
KW   apoptosis; immunosuppression; antiinflammatory; immunostimulation;
KW   probe; primer; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO200034294-A2.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;   99WO-US29400.
XX
PR   11-DEC-1998;   98US-0111826.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Bowen MA,  Siemers N;
XX
DR   WPI; 2000-423364/36.
XX
PT   Novel tumor necrosis factor receptor homologue-1 useful as a target for
PT   immunosuppressive, antiinflammatory and/or immunostimulatory drug
PT   development
XX
PS   Disclosure; Fig 1b; 42pp; English.
XX
CC   The tumour necrosis factor receptor homologue TRH1 can be used for
CC   treating a mammal e.g. a human, at risk for a disorder characterized
CC   by an aberrant or unwanted level or biological activity of TRH1,
CC   e.g. rheumatoid arthritis and transplant rejection. TRH1 may also be
CC   useful to leach out or block a ligand which is found to bind to the
CC   TRH1. TRH1 may be used in various drug screening techniques and to
CC   identify fragments and analogs of a protein or peptide (agonist or
CC   antagonist) which bind to TRH1. The TRH1 protein plays a role in
CC   cellular function, cell activation, proliferation, differentiation,
CC   and apoptosis. The interaction between the novel TNFr protein of the
CC   present invention and intracellular signaling molecules and/or its
CC   potential co-receptor may serve as a novel target for
CC   immunosuppressive, antiinflammatory and/or immunostimulatory drug
CC   development. Gene constructs can also be used as part of a gene
CC   therapy protocol to deliver nucleic acids encoding the TRH1, or an
CC   agonist or antagonist form of a TRH1 protein or peptide. Antibody
CC   directed against TRH1 can be used to reject TRH1 in tissues
CC   and cells. They can also be used to make targeted antibody that
CC   destroy TRH1 expressing cells. Fragments of the TRH1 gene can be
CC   used as diagnostic probes or as PCR primers. Fragments of the full
CC   length gene may be used as hybridization probes for a cDNA library to
CC   isolate the full length gene and to isolate other genes which have a
CC   high sequence similarity. The probes may be used to identify a cDNA
CC   clone corresponding to a full length transcript and a genomic clone
CC   or clones that contain the complete gene including regulatory and
CC   promoter regions, exons, and introns.
XX
SQ   Sequence 3236 BP; 810 A; 808 C; 751 G; 867 T; 0 other;


alignment_scores:
            Quality: 2091.00              Length:    405
              Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAA47396    ..

 Align seg 1/1  to: AAA47396  from: 1  to: 3236

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     765 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 814

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     815 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 864
```

```
  34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 865 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 914

  51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 915 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 964

  67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 965 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1014

  84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1015 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1064

 101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1065 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1114

 117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1115 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1164

 134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1165 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1214

 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1215 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1264

 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1265 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1314

 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1315 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1364

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1365 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1414

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1415 CATCCGGGGCCCCGAGGCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1464

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1465 AGCACCGGAGAAACGATGTTCTGGAGAAGATTCGTGGGCTGATGGAAGAC 1514

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1515 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1564

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1565 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1614

 284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1615 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1664

 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1665 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1714

 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1715 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1764

 334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1765 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 1814

 351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1815 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 1864

 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1865 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 1914

 384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1915 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 1964

 401 LeuProAspLeuLeu 405
     |||||||||||||||
1965 CTTCCTGACCTGCTG 1979
```

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA47395

seq_documentation_block:

```
ID   AAA47395 standard; cDNA; 3277 BP.
XX
AC   AAA47395;
XX
DT   20-OCT-2000  (first entry)
XX
DE   Tumour necrosis factor receptor homologue TRH1 (Clone 2733717).
XX
KW   Tumour necrosis factor receptor homologue; TRH1; TNF; arthritis;
KW   transplant rejection; activation; proliferation; differentiation;
KW   apoptosis; immunosuppression; antiinflammatory; immunostimulation;
KW   probe; primer; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             52..2019
FT                   /*tag=  a
FT                   /product= Tumour necrosis factor receptor homologue
XX
PN   WO200034294-A2.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;   99WO-US29400.
XX
PR   11-DEC-1998;   98US-0111826.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Bowen MA,  Siemers N;
XX
DR   WPI; 2000-423364/36.
DR   P-PSDB; AAB01349.
XX
PT   Novel tumor necrosis factor receptor homologue-1 useful as a target for
PT   immunosuppressive, antiinflammatory and/or immunostimulatory drug
PT   development
XX
PS   Claim 2; Fig 2a-2d; 42pp; English.
XX
CC   The tumour necrosis factor receptor homologue TRH1 can be used for
CC   treating a mammal e.g. a human, at risk for a disorder characterized
CC   by an aberrant or unwanted level or biological activity of TRH1,
CC   e.g. rheumatoid arthritis and transplant rejection. TRH1 may also be
CC   useful to leach out or block a ligand which is found to bind to the
CC   TRH1. TRH1 may be used in various drug screening techniques and to
CC   identify fragments and analogs of a protein or peptide (agonist or
CC   antagonist) which bind to TRH1. The TRH1 protein plays a role in
CC   cellular function, cell activation, proliferation, differentiation,
CC   and apoptosis. The interaction between the novel TNFr protein of the
```

```
CC   present invention and intracellular signaling molecules and/or its
CC   potential co-receptor may serve as a novel target for
CC   immunosuppressive, antiinflammatory and/or immunostimulatory drug
CC   development. Gene constructs can also be used as part of a gene
CC   therapy protocol to deliver nucleic acids encoding the TRH1, or an
CC   agonist or antagonist form of a TRH1 protein or peptide. Antibody
CC   directed against TRH1 can be used to reject TRH1 in tissues
CC   and cells. They can also be used to make targeted antibody that
CC   destroy TRH1 expressing cells. Fragments of the TRH1 gene can be
CC   used as diagnostic probes or as PCR primers. Fragments of the full
CC   length gene may be used as hybridization probes for a cDNA library to
CC   isolate the full length gene and to isolate other genes which have a
CC   high sequence similarity. The probes may be used to identify a cDNA
CC   clone corresponding to a full length transcript and a genomic clone
CC   or clones that contain the complete gene including regulatory and
CC   promoter regions, exons, and introns.
XX
SQ   Sequence 3277 BP; 812 A; 825 C; 765 G; 875 T; 0 other;


alignment_scores:
            Quality: 2091.00              Length:    405
              Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAA47395   ..

 Align seg 1/1  to: AAA47395  from: 1  to: 3277

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     802 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 851

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     852 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 901

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     902 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 951

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     952 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1001

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1002 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1051

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1052 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1101

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1102 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1151

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1152 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1201

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1202 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1251

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1252 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1301

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1302 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1351

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1352 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1401

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1402 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1451

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1452 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1501

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1502 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1551

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1552 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1601

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1602 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1651

     284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1652 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1701

     301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1702 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1751

     317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1752 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1801

     334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1802 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 1851

     351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1852 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 1901

     367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1902 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 1951

     384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1952 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 2001

     401 LeuProAspLeuLeu 405
         |||||||||||||||
    2002 CTTCCTGACCTGCTG 2016

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25259

seq_documentation_block:
ID   AAX25259 standard; cDNA; 3440 BP.
XX
AC   AAX25259;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Human full-length tumour necrosis factor receptor ZTNFR-6 cDNA.
XX
KW   ZTNFR-6; tumour necrosis factor receptor-6; human;
KW   cell maturation; bone cell regulation; ss.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             284..2179
FT                   /*tag= a
FT                   /product= "membrane-bound ZTNFR-6"
FT   misc_feature    284..2225
FT                   /*tag= b
FT                   /note= "this region of the sequence is
FT                         specifically claimed in Claim 10(a)"
FT   sig_peptide     284..334
FT                   /*tag= c
FT   mat_peptide     335..2176
FT                   /*tag= d
XX
PN   WO9911790-A1.
XX
PD   11-MAR-1999.
XX
PF   03-SEP-1998;   98WO-US18364.
XX
PR   04-SEP-1997;   97US-0923725.
PR   04-SEP-1997;   97US-0057608.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Farrah TM,  Gross JA,  Matthews SM;
XX
DR   WPI; 1999-205190/17.
DR   P-PSDB; AAY05678.
XX
PT   New secreted or membrane bound tumor necrosis factor receptor
PT   ZTNFR-6  - useful for detecting a genetic abnormality in a patient
XX
PS   Claim 10a; Page 108-113; 145pp; English.
XX
CC   This is the DNA sequence of an isolated polynucleotide that codes
CC   for a novel human full-length membrane-bound tumour necrosis factor
CC   receptor, designated ZTNFR-6 (see AAY05678).  The polynucleotide was
CC   initially identified by querying an EST database for sequences
CC   homologous to conserved motifs within the TNFR family.  A second
CC   clone (see AAX25260) encodes truncated, soluble ZTNFR-6 (see AAY05679).
CC   The full-length clone includes a 40 bp exon (see AAX25262) that puts
CC   the sequence in frame for the transmembrane and cytoplasmic domains
CC   of membrane-bound ZTNFR-6.  The human ZTNFR-6 gene was localised to
CC   6p21.1.  The isolated polynucleotides can be utilised in the
CC   recombinant production of ZTNFR-6 polypeptides, and also in a
CC   claimed method for detecting a genetic abnormality in a patient.
CC   ZTNFR-6 polypeptides are useful in methods that promote cellular
CC   maturation and bone cell regulation.
XX
SQ   Sequence 3440 BP; 827 A; 876 C; 833 G; 904 T; 0 other;


alignment_scores:
           Quality: 2091.00              Length:    405
             Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAX25259    ..

 Align seg 1/1  to: AAX25259  from: 1  to: 3440

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     962 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1011

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1012 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1061

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1062 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1111

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1112 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1161

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1162 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1211

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1212 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1261

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1262 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1311

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1312 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1361

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1362 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1411

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1412 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1461

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1462 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1511

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1512 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1561

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1562 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1611

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1612 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1661

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1662 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1711

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1712 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1761

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1762 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1811

     284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1812 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1861

     301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1862 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1911

     317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1912 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1961

     334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
         |||||||||||||||||||||||||||||||||||||||||||||||||||
    1962 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2011
```

```
     351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2012 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2061

     367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2062 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2111

     384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    2112 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 2161

     401 LeuProAspLeuLeu 405
         ||||||||||||||||
    2162 CTTCCTGACCTGCTG 2176
```

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV99927

seq_documentation_block:

```
ID   AAV99927 standard; DNA; 3474 BP.
XX
AC   AAV99927;
XX
DT   10-MAY-1999  (first entry)
XX
DE   Novel human tumor necrosis factor receptor TR9.
XX
KW   Human; tumour necrosis factor receptor; TNFR; TR9 receptor; cancer;
KW   apoptosis; agonist; inhibition; autoimmune disorder; viral infection;
KW   inflammation; antagonist; AIDS; neurodegenerative disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             247..2214
FT                   /*tag=  a
FT                   /product= "human TR9 receptor"
FT   sig_peptide     247..366
FT                   /*tag=  b
FT   mat_peptide     367..2211
FT                   /*tag=  c
XX
PN   WO9856892-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-US11932.
XX
PR   11-JUN-1997;   97US-0052991.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Fan P,  Gentz RL,  Ni J,  Yu G;
XX
DR   WPI; 1999-060325/05.
DR   P-PSDB; AAW81059.
XX
PT   New isolated tumour necrosis factor-like receptor, TR9 - used to
PT   develop products for treating e.g. cancers, autoimmune disorders,
PT   viral infections, inflammation, graft rejection, neurodegenerative
PT   disorders or ischaemic injury
XX
PS   Claim 2; Fig 1; 134pp; English.
XX
CC   This is the nucleotide sequence of the human tumour necrosis factor
CC   receptor (TNFR), TR9 receptor, used in the method of the invention
CC   to develop products to treat disorders such as cancers. The novel
CC   TNFR, TR9, can be used to identify agents for modifying apoptosis.
CC   Agonists can be used to treat diseases associated with increased
CC   cell survival, or the inhibition of apoptosis, including cancers
CC   (e.g. follicular lymphomas, carcinomas with p53 mutations, and
CC   hormone-dependent tumours, such as breast cancer, prostate cancer,
CC   Kaposi's sarcoma and ovarian cancer), autoimmune disorders (e.g.
CC   systemic lupus erythematosus and immune-related glomerulonephritis
CC   rheumatoid arthritis), viral infections (e.g. herpes viruses, pox
CC   viruses and adenoviruses), inflammation, graft vs host disease,
CC   acute graft rejection and chronic graft rejection. Antagonists
CC   can be used to treat diseases associated with increased apoptosis
CC   including AIDS, neurodegenerative disorders (e.g. Alzheimer's
CC   disease, Parkinson's disease, amyotrophic lateral sclerosis,
CC   retinitis pigmentosa, cerebellar degeneration), myelodysplstic
CC   syndromes (e.g. aplastic anemia), ischemic injury (e.g. that
CC   caused by myocardial infarction, stroke and reperfusion injury),
CC   toxin-induced liver disease (e.g. that caused by alcohol), septic
CC   shock, cachexia, anorexia, inflammatory diseases and stress response
CC   related diseases, such as inflammatory bowel disease, rheumatoid
CC   arthritis, osteoarthritis, psoriasis and septicemia. The products
CC   can also be used for detection and diagnosis.
XX
SQ   Sequence 3474 BP; 832 A; 895 C; 844 G; 903 T; 0 other;
```

alignment_scores:

```
            Quality: 2091.00              Length:    405
              Ratio:  5.163                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:

```
 US-09-733-956-2 x AAV99927   ..

 Align seg 1/1  to: AAV99927  from: 1  to: 3474

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     997 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1046

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1047 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1096

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1097 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1146

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1147 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1196

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1197 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1246

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1247 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1296

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1297 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1346

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1347 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1396

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1397 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1446

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1447 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1496

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
    1497 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1546
```

```
 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1547 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1596

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1597 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1646

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1647 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1696

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1697 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1746

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1747 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1796

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1797 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1846

 284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1847 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1896

 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1897 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1946

 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1947 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1996

 334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1997 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2046

 351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2047 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2096

 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2097 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2146

 384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2147 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 2196

 401 LeuProAspLeuLeu 405
     |||||||||||||||
2197 CTTCCTGACCTGCTG 2211
```

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA99658

seq_documentation_block:

```
ID   AAA99658 standard; cDNA; 3474 BP.
XX
AC   AAA99658;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Human tumour necrosis factor receptor TR9 cDNA.
XX
KW   Human; tumour necrosis factor; TR9 receptor; immunosuppressive;
KW   antiinflammatory; cardiant; antiasthmatic; antidiabetic; antiallergic;
KW   antiathritic; antirheumatic; anti-HIV; anticonvulsant; cytostatic;
KW   neuroprotective; gene therapy; Death Domain Containing Receptor 6;
KW   common variable immunodeficiency; X-linked agammaglobulinaemia;
KW   severe combined immunodeficiency; Wiskott-Aldrich syndrome;
KW   autoimmune disease; rheumatoid arthritis; allergic encephalomyelitis;
KW   multiple sclerosis; diabetes mellitus; asthma; epilepsy; cancer;
KW   cardiovascular disease; neurological disease;
KW   protein coordinate data; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             247..2214
FT                   /*tag= a
FT                   /product= "TR9 receptor"
FT   sig_peptide     247..366
FT                   /*tag= b
FT   mat_peptide     367..2214
FT                   /*tag= c
XX
PN   WO200056862-A1.
XX
PD   28-SEP-2000.
XX
PF   16-MAR-2000; 2000WO-US06831.
XX
PR   24-MAR-1999;   99US-0126019.
PR   14-MAY-1999;   99US-0134220.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J,  Gentz RL,  Yu G,  Fan P;
XX
DR   WPI; 2000-594575/56.
DR   P-PSDB; AAB26981.
XX
PT   Nucleic acid molecule encoding a human tumor necrosis factor receptor,
PT   known as TR9, useful for treating, preventing and diagnosing severe
PT   combined immunodeficiency, autoimmune diseases, HIV infection, epilepsy
PT   and cancer -
XX
PS   Claim 2; Fig 1; 220pp; English.
XX
CC   The present sequence encodes a novel human tumour necrosis factor
CC   receptor, designated TR9. The TR9 receptor is also known as Death Domain
CC   Containing Receptor 6. TR9 polypeptides, polynucleotides or agonists are
CC   useful for treating, preventing or diagnosing common variable
CC   immunodeficiency, X-linked agammaglobulinaemia, severe combined
CC   immunodeficiency and Wiskott-Aldrich syndrome, autoimmune diseases (such
CC   as rheumatoid arthritis, allergic encephalomyelitis, multiple sclerosis,
CC   diabetes mellitus and asthma), HIV infection, epilepsy, cancer,
CC   cardiovascular diseases and other neurological diseases.
XX
SQ   Sequence 3474 BP; 832 A; 895 C; 844 G; 903 T; 0 other;
```

alignment_scores:

```
           Quality: 2091.00              Length:    405
             Ratio:  5.163                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:

```
 US-09-733-956-2 x AAA99658   ..

 Align seg 1/1  to: AAA99658  from: 1  to: 3474

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     997 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1046

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1047 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1096

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1097 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1146
```

```
  51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1147 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1196

  67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1197 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1246

  84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1247 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1296

 101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1297 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1346

 117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1347 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1396

 134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1397 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1446

 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1447 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1496

 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1497 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1546

 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1547 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1596

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1597 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1646

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1647 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1696

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1697 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1746

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1747 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1796

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1797 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1846

 284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1847 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1896

 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1897 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1946

 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1947 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1996

 334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1997 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2046

 351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
     |||||||||||||||||||||||||||||||||||||||||||||||||||
2047 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2096

 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
     |||||||||||||||||||||||||||||||||||||||||||||||||||
2097 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2146

 384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
     |||||||||||||||||||||||||||||||||||||||||||||||||||
2147 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 2196

 401 LeuProAspLeuLeu 405
     ||||||||||||||
2197 CTTCCTGACCTGCTG 2211
```

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ33945

seq_documentation_block:

```
ID   AAZ33945 standard; cDNA; 3534 BP.
XX
AC   AAZ33945;
XX
DT   07-DEC-1999  (first entry)
XX
DE   Human PRO868 nucleotide sequence.
XX
KW   Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW   probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW   secreted protein; transmembrane protein; ss.
XX
OS   Homo sapiens.
XX
PN   WO9946281-A2.
XX
PD   16-SEP-1999.
XX
PF   08-MAR-1999;   99WO-US05028.
XX
PR   10-MAR-1998;   98US-0077450.
PR   11-MAR-1998;   98US-0077632.
PR   11-MAR-1998;   98US-0077641.
PR   11-MAR-1998;   98US-0077649.
PR   12-MAR-1998;   98US-0077791.
PR   13-MAR-1998;   98US-0078004.
PR   17-MAR-1998;   98US-0040220.
PR   20-MAR-1998;   98US-0078886.
PR   20-MAR-1998;   98US-0078910.
PR   20-MAR-1998;   98US-0078936.
PR   20-MAR-1998;   98US-0078939.
PR   25-MAR-1998;   98US-0079294.
PR   26-MAR-1998;   98US-0079656.
PR   27-MAR-1998;   98US-0079663.
PR   27-MAR-1998;   98US-0079664.
PR   27-MAR-1998;   98US-0079689.
PR   27-MAR-1998;   98US-0079728.
PR   27-MAR-1998;   98US-0079786.
PR   30-MAR-1998;   98US-0079920.
PR   30-MAR-1998;   98US-0079923.
PR   31-MAR-1998;   98US-0080105.
PR   31-MAR-1998;   98US-0080107.
PR   31-MAR-1998;   98US-0080165.
PR   31-MAR-1998;   98US-0080194.
PR   01-APR-1998;   98US-0080327.
PR   01-APR-1998;   98US-0080328.
PR   01-APR-1998;   98US-0080333.
PR   01-APR-1998;   98US-0080334.
PR   08-APR-1998;   98US-0081049.
PR   08-APR-1998;   98US-0081070.
PR   08-APR-1998;   98US-0081071.
PR   09-APR-1998;   98US-0081195.
PR   09-APR-1998;   98US-0081203.
PR   09-APR-1998;   98US-0081229.
```

```
PR   15-APR-1998;   98US-0081817.
PR   15-APR-1998;   98US-0081838.
PR   15-APR-1998;   98US-0081952.
PR   15-APR-1998;   98US-0081955.
PR   21-APR-1998;   98US-0082568.
PR   21-APR-1998;   98US-0082569.
PR   22-APR-1998;   98US-0082700.
PR   22-APR-1998;   98US-0082704.
PR   22-APR-1998;   98US-0082804.
PR   23-APR-1998;   98US-0082767.
PR   23-APR-1998;   98US-0082796.
PR   27-APR-1998;   98US-0083336.
PR   28-APR-1998;   98US-0083322.
PR   29-APR-1998;   98US-0083392.
PR   29-APR-1998;   98US-0083495.
PR   29-APR-1998;   98US-0083496.
PR   29-APR-1998;   98US-0083499.
PR   29-APR-1998;   98US-0083500.
PR   29-APR-1998;   98US-0083545.
PR   29-APR-1998;   98US-0083554.
PR   29-APR-1998;   98US-0083558.
PR   29-APR-1998;   98US-0083559.
PR   30-APR-1998;   98US-0083742.
PR   05-MAY-1998;   98US-0084366.
PR   06-MAY-1998;   98US-0084414.
PR   06-MAY-1998;   98US-0084441.
PR   07-MAY-1998;   98US-0084598.
PR   07-MAY-1998;   98US-0084600.
PR   07-MAY-1998;   98US-0084627.
PR   07-MAY-1998;   98US-0084637.
PR   07-MAY-1998;   98US-0084639.
PR   07-MAY-1998;   98US-0084640.
PR   07-MAY-1998;   98US-0084643.
PR   13-MAY-1998;   98US-0085323.
PR   13-MAY-1998;   98US-0085338.
PR   13-MAY-1998;   98US-0085339.
PR   15-MAY-1998;   98US-0085573.
PR   15-MAY-1998;   98US-0085579.
PR   15-MAY-1998;   98US-0085580.
PR   15-MAY-1998;   98US-0085582.
PR   15-MAY-1998;   98US-0085689.
PR   15-MAY-1998;   98US-0085697.
PR   15-MAY-1998;   98US-0085700.
PR   15-MAY-1998;   98US-0085704.
PR   18-MAY-1998;   98US-0086023.
PR   22-MAY-1998;   98US-0086392.
PR   22-MAY-1998;   98US-0086414.
PR   22-MAY-1998;   98US-0086430.
PR   22-MAY-1998;   98US-0086486.
PR   28-MAY-1998;   98US-0087098.
PR   28-MAY-1998;   98US-0087106.
PR   28-MAY-1998;   98US-0087208.
PR   30-JUL-1998;   98US-0094651.
PR   11-SEP-1998;   98US-0100038.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wood WI,  Goddard A,  Gurney A,  Yuan J,  Baker KP,  Chen J;
XX
DR   WPI; 1999-551358/46.
DR   P-PSDB; AAY41693.
XX
PT   New secreted and transmembrane polypeptides and their polynucleotides,
PT   useful for treating blood coagulation disorders, cancers and cellular
PT   adhesion disorders -
XX
PS   Claim 2; Fig 25; 530pp; English.
XX
CC   The present invention describes secreted and transmembrane polypeptides
CC   and their polynucleotides. The nucleotide sequences are useful as
CC   sources of probes, primers, for chromosome mapping, and for generation
CC   of antisense sequences. They can also be used to create transgenic
CC   animals. The proteins can be used to treat a variety of diseases and
CC   disorders, depending on their function. Diseases that may be treated
CC   include blood coagulation disorders, cancers and cellular adhesion
CC   disorders. They may also be used to raise antibodies. AAZ33891 to
CC   AAZ34338, and AAY41685 to AAY41774 represent polynucleotide and
CC   polypeptide sequence given in the exemplification of the present
CC   invention.
XX
SQ   Sequence 3534 BP; 829 A; 927 C; 858 G; 920 T; 0 other;


alignment_scores:
           Quality: 2091.00              Length:    405
             Ratio:  5.163                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAZ33945   ..

 Align seg 1/1  to: AAZ33945  from: 1  to: 3534

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1075 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1124

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1125 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1174

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1175 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1224

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1225 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1274

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1275 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1324

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1325 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1374

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1375 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1424

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1425 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1474

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1475 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1524

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1525 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1574

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1575 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1624

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1625 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1674

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1675 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1724

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  1725 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1774

   234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1775 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1824

   251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1825 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1874

   267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1875 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1924

   284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1925 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1974

   301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1975 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 2024

   317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  2025 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 2074

   334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  2075 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2124

   351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  2125 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2174

   367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  2175 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2224

   384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  2225 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 2274

   401 LeuProAspLeuLeu 405
       |||||||||||||||
  2275 CTTCCTGACCTGCTG 2289

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC78474

seq_documentation_block:
ID   AAC78474 standard; cDNA; 3534 BP.
XX
AC   AAC78474;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human PRO868 (UNQ437) nucleotide sequence SEQ ID NO:63.
XX
KW   Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;
KW   expressed sequence tag; detection; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200053756-A2.
XX
PD   14-SEP-2000.
XX
PF   18-FEB-2000; 2000WO-US04341.
XX
PR   08-MAR-1999;   99WO-US05028.
PR   12-MAR-1999;   99US-0123957.
PR   29-MAR-1999;   99US-0126773.
PR   21-APR-1999;   99US-0130232.
PR   28-APR-1999;   99US-0131445.
PR   14-MAY-1999;   99US-0134287.
PR   23-JUN-1999;   99US-0141037.
PR   26-JUL-1999;   99US-0145698.
PR   29-OCT-1999;   99US-0162506.
PR   30-NOV-1999;   99WO-US28313.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28565.
PR   16-DEC-1999;   99WO-US30095.
PR   30-DEC-1999;   99WO-US31243.
PR   30-DEC-1999;   99WO-US31274.
PR   05-JAN-2000; 2000WO-US00219.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Baker KP,  Botstein D,  Desnoyers L,  Eaton DL;
PI   Ferrara N,  Filvaroff E,  Fong S,  Gao W,  Gerber H,  Gerritsen ME;
PI   Goddard A,  Godowski PJ,  Grimaldi CJ,  Gurney AL,  Hillan KJ;
PI   Kljavin IJ,  Kuo SS,  Napier MA,  Pan J,  Paoni NF,  Roy MA;
PI   Shelton DL,  Stewart TA,  Tumas D,  Williams PM,  Wood WI;
XX
DR   WPI; 2000-611443/58.
DR   P-PSDB; AAB44249.
XX
PT   Novel PRO polypeptides and polynucleotides used in detection methods,
PT   to target bioactive molecules to specific cells, and to modulate
PT   cellular activities -
XX
PS   Claim 2; Fig 25; 636pp; English.
XX
CC   AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC   sequence tag) sequences which encode secreted or transmembrane PRO
CC   polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC   activity. The polynucleotides and polypeptides can be used for detecting
CC   the presence of PRO polypeptides in samples, for linking bioactive
CC   molecules to cells and for modulating biological activities of cells,
CC   using the polypeptides for specific targeting. The polypeptide targeting
CC   can be used to kill the target cells, e.g. for the treatment of cancers.
CC   The polypeptide pairs provide specific targeting of bioactive molecules
CC   to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC   the isolation of the PRO polynucleotide sequences.
XX
SQ   Sequence 3534 BP; 829 A; 927 C; 858 G; 920 T; 0 other;


alignment_scores:
            Quality: 2091.00              Length:    405
              Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAC78474   ..

 Align seg 1/1  to: AAC78474  from: 1  to: 3534

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1075 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1124

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1125 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1174

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1175 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1224

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1225 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1274
```

```
  67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1275 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1324

  84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1325 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1374

 101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1375 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1424

 117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1425 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1474

 134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1475 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1524

 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1525 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1574

 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1575 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1624

 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1625 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1674

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1675 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1724

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1725 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1774

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1775 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1824

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1825 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1874

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1875 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1924

 284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1925 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1974

 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1975 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 2024

 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2025 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 2074

 334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2075 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2124

 351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2125 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2174

 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2175 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2224

 384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2225 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 2274

 401 LeuProAspLeuLeu 405
     ||||||||||||||||
2275 CTTCCTGACCTGCTG 2289
```

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH14374

seq_documentation_block:

ID AAH14374 standard; cDNA; 2636 BP.
XX
AC AAH14374;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:11786.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 11786; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification

```
CC   of the present invention.
XX
SQ   Sequence 2636 BP; 569 A; 762 C; 722 G; 583 T; 0 other;

alignment_scores:
            Quality: 2087.00          Length:    405
              Ratio:  5.153            Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.753

alignment_block:
 US-09-733-956-2 x AAH14374    ..

 Align seg 1/1  to: AAH14374  from: 1  to: 2636

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1145 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1194

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1195 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1244

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1245 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1294

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1295 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1344

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1345 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1394

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1395 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1444

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1445 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1494

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1495 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1544

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1545 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1594

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1595 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1644

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1645 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1694

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1695 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1744

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1745 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1794

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1795 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1844

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
    1845 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1894

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1895 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1944

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1945 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1994

     284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1995 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 2044

     301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2045 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 2094

     317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2095 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 2144

     334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2145 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2194

     351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2195 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2244

     367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2245 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2294

     384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
         ||||||||||||||||||||||||||||||:::||||||||||||||||||
    2295 TCAAGAGCCAGGAAGCCAGCCAGACCTTCCTGGACTCTGTTTATAGCCAT 2344

     401 LeuProAspLeuLeu 405
         ||||||||||||||||
    2345 CTTCCTGACCTGCTG 2359

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25320

seq_documentation_block:
ID   AAX25320 standard; cDNA; 3331 BP.
XX
AC   AAX25320;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Mouse TNF receptor TRL cDNA.
XX
KW   TRL; TNF receptor-like; tumour necrosis factor receptor; mouse;
KW   signal transduction; cell differentiation; prostate cancer;
KW   inflammation; arthritis; diabetes; insulin resistance; diagnosis;
KW   therapy; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             344..2065
FT                   /*tag=  a
XX
PN   WO9915663-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;    98WO-US20219.
XX
PR   17-MAR-1998;    98US-0042785.
PR   26-SEP-1997;    97US-0938896.
XX
```

```
PA   (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI   Busfield SJ;
XX
DR   WPI; 1999-254712/21.
DR   P-PSDB; AAY05695.
XX
PT   Nucleic acid encoding tumor necrosis factor receptor-like protein
XX
PS   Claim 1b; Page 107-113; 170pp; English.
XX
CC   This cDNA sequence codes for a novel member of the murine tumour
CC   necrosis factor receptor superfamily (see AAY05695), termed TRL (TNF
CC   receptor-like). The cDNA was isolated from a primary murine
CC   megakaryocyte cDNA library, and was used to identify human TRL I
CC   and TRL II clones (see AAX25321 and AAX25322). The invention provides
CC   full-length TRL proteins, TRL fusion proteins, antigenic peptides
CC   and anti-TRL antibodies, as well as TRL nucleic acids, recombinant
CC   expression vectors, host cells and non-human transgenic animals.
CC   TRL are regulators or modulators of cellular signal transduction,
CC   cellular proliferation or differentiation, cell survival and
CC   apoptosis, immune system cells, and cells involved in insulin
CC   resistance or the diabetic response. They can be used to isolate
CC   cognate ligands, to modulate ligand/TRL interactions, to screen
CC   for potential modulators, and to treat conditions associated with
CC   abnormal TRL levels. TRL nucleic acids are used for recombinant
CC   production of TRL, to detect mutations, to identify TRL homologues,
CC   as antisense, triplex-forming or ribozyme therapeutics, for
CC   chromosome mapping and for forensic identification of individuals.
CC   Antibodies, antisense sequences, TRL protein mutants etc. that
CC   modulate activity of TRL are used to treat or prevent conditions
CC   associated with aberrant TRL protein or expression, e.g. cell
CC   proliferation and differentiation disorders (particularly cancer,
CC   especially of the prostate), inflammation (e.g. arthritis) and
CC   metabolic disease (e.g. diabetes and insulin resistance).
XX
SQ   Sequence 3331 BP; 825 A; 891 C; 739 G; 876 T; 0 other;

alignment_scores:
           Quality: 1914.00          Length:    406
             Ratio:  4.883            Gaps:      2
 Percent Similarity:  96.552   Percent Identity: 92.857

alignment_block:
 US-09-733-956-2 x AAX25320   ..

 Align seg 1/1  to: AAX25320  from: 1  to: 3331

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||:::|||||||||:::||||||||||||   ||||||
     848 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 897

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         |||:::|||:::|||||||||||||||||||||||||||||:::   |
     898 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 947

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||   :::|||:::||||||||||||         |||:::|||
     948 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 997

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
         ||||||:::|||||||||||||||||||||||||||||||   |||||
     998 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 1047

      66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
         ||||||||||   ||||||||||||||||   ||||||:::|||||||
    1048 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 1094

      83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
         ||||||||||||||||||   ||||||||||||||||||||||||||||
    1095 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 1144

     100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
    1145 TTGCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGT 1194

     116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1195 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 1244

     133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
         ||||||||||||||||||||||||||||||||||||||||||:::|||
    1245 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 1294

     150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
         |||||||||||||||||||||||||||||||    ||||||||||||||
    1295 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 1344

     166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1345 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACA 1394

     183 leTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1395 TCTATCAGTTTCTTTGCAACGCCAGTGAGAGGGAGGTGGCGGCCTTCTCC 1444

     200 AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1445 AATGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCAGCACTG 1494

     216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1495 GACCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGC 1544

     233 rgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1545 GCCAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAA 1594

     250 AspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSe 266
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1595 GACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAGCCCCAG 1644

     266 rProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnS 283
         ||||||||||||||||||||||||||||||||:::|||||||||||||||
    1645 TCCGCTGAGCCCGAGCCCCATCCCCAGTCCTAACGTGAAACTTGAGAATT 1694

     283 erAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
         ||:::|||||||||||||||||||||||||   ||||||||||||   |||
    1695 CCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAGTGCTTC 1744

     300 PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGl 316
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1745 TTCGTGGACGAGTCAGAGCCCCTTCTGCGTTGCGACTCCACATCCAGTGG 1794

     316 ySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysA 333
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1795 CTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAAGAAAAGAAGG 1844

     333 spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1845 ACACAGTGTTGCGGCAGGTCCGCCTGGACCCCTGTGACTTGCAGCCCATC 1894

     350 PheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGl 366
         ||||||||||||||||||:::||||||||||||||||||||||||||||
    1895 TTTGATGACATGCTGCATATCCTGAACCCCGAGGAGCTGCGGGTGATTGA 1944

     366 uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleG 383
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1945 AGAGATTCCCCAGGCTGAGGACAAACTGGACCGCCTCTTCGAGATCATTG 1994

     383 lyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSer 399
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1995 GGGTCAAGAGCCAAGAAGCCAGCCAGACCCTCTTGGACTCTGTGTACAGT 2044

     400 HisLeuProAspLeuLeu 405
```

```
         ||||||||||||||||||
    2045 CATCTTCCTGACCTATTG 2062

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25264

seq_documentation_block:
ID   AAX25264 standard; DNA; 1893 BP.
XX
AC   AAX25264;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Degenerate nucleotide sequence encoding ZTNFR-6.
XX
KW   ZTNFR-6; tumour necrosis factor receptor-6; human;
KW   cell maturation; bone cell regulation; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9911790-A1.
XX
PD   11-MAR-1999.
XX
PF   03-SEP-1998;   98WO-US18364.
XX
PR   04-SEP-1997;   97US-0923725.
PR   04-SEP-1997;   97US-0057608.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Farrah TM,  Gross JA,  Matthews SM;
XX
DR   WPI; 1999-205190/17.
DR   P-PSDB; AAY05678.
XX
PT   New secreted or membrane bound tumor necrosis factor receptor
PT   ZTNFR-6  - useful for detecting a genetic abnormality in a patient
XX
PS   Claim 10c; Page 129-130; 145pp; English.
XX
CC   This degenerate nucleotide sequence encodes a novel human
CC   membrane-bound tumour necrosis factor receptor, designated ZTNFR-6
CC   (see AAY05678).  It includes all possible codons for the given
CC   amino acid sequence of ZTNFR-6.  The degenerate sequence can
CC   serve as a template for optimising expression of ZTNFR-6
CC   polynucleotides in various cell types and species.  Sequences
CC   containing preferential codons can be tested and optimised for
CC   expression in various species, and tested for functionality.
CC   ZTNFR-6 is a novel membrane of the TNFR family.  ZTNFR-6
CC   polypeptides can be used to promote cellular maturation and
CC   bone cell regulation.  Polynucleotides can be used in the
CC   recombinant production of ZTNFR-6 polypeptides, and in a method for
CC   detecting a genetic abnormality in a patient.
XX
SQ   Sequence 1893 BP; 338 A; 224 C; 293 G; 211 T; 827 other;


alignment_scores:
           Quality: 1567.00            Length:    403
             Ratio:  4.650               Gaps:      0
 Percent Similarity:  83.623   Percent Identity: 73.449

alignment_block:
 US-09-733-956-2 x AAX25264   ..

 Align seg 1/1  to: AAX25264  from: 1  to: 1893

      1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
        ||||||:::||||||:::|||::::::|||:::|||    ||||||||
    679 ATGAAYWSNACNGARWSNAAYWSNWSNGCNWSNGTNMGNCCNAARGTNYT 728

     17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
        ::::::|||||||||||||||||||||||||||||||||::::::|||    |
    729 NWSNWSNATHCARGARGGNACNGTNCCNGAYAAYACNWSNWSNGCNMGNG 778

     34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
        ||||||||||||||||||||||    ||||||    ||||||||||||||
    779 GNAARGARGAYGTNAAYAARACNYTNCCNAAYYTNCARGTNGTNAAYCAY 828

     51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
        |||||||||||||||||    ||||||        |||:::|||||
    829 CARCARGGNCCNCAYCAYMGNCAYATHYTNAARYTNYTNCCNWSNATGGA 878

     67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
        ||||||||||||||||||::::::||||||||||||||||||||    |
    879 RGCNACNGGNGGNGARAARWSNWSNACNCCNATHAARGGNCCNAARMGNG 928

     84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
        ||||||||    ||||||    ||||||||||||||||||||||||||
    929 GNCAYCCNMGNCARAAYYTNCAYAARCAYTTYGAYATHAAYGARCAYYTN 978

    101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
        ||||||||||||||    |||          |||    |||||||||||||
    979 CCNTGGATGATHGTNYTNTTYYTNYTNYTNGTNYTNGTNGTNATHGTNGT 1028

    117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
        ||||:::|||    |||::::::    |||    ||||||||||||    ||||
   1029 NTGYWSNATHMGNAARWSNWSNMGNACNYTNAARAARGGNCCNMGNCARG 1078

    134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
        |||||:::|||||||||||||||||||||    |||||:::|||||||||
   1079 AYCCNWSNGCNATHGTNGARAARGCNGGNYTNAARAARWSNATGACNCCN 1128

    151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
        |||||||||    ||||||||||||||||||||||||||||||||||||||
   1129 ACNCARAAYMGNGARAARTGGATHTAYTAYTGYAAYGGNCAYGGNATHGA 1178

    167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
        ||||    |||    ||||||||||||||||||:::||||||||||||||||
   1179 YATHYTNAARYTNGTNGCNGCNCARGTNGGNWSNCARTGGAARGAYATHT 1228

    184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
        ||||||||    |||||||||:::|||    |||||||||||||||:::|||
   1229 AYCARTTYYTNTGYAAYGCNWSNGARMGNGARGTNGCNGCNTTYWSNAAY 1278

    201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
        |||||||||||||||||||||    ||||||||||||    |||||||||||
   1279 GGNTAYACNGCNGAYCAYGARMGNGCNTAYGCNGCNYTNCARCAYTGGAC 1328

    217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
        ||||    |||||||||||||:::    ||||||    |||:::|||        |
   1329 NATHMGNGGNCCNGARGCNWSNYTNGCNCARYTNATHWSNGCNYTNMGNC 1378

    234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
        |||||        ||||||||||||||||||||||    |||    |||||||||
   1379 ARCAYMGNMGNAAYGAYGTNGTNGARAARATHMGNGGNYTNATGGARGAY 1428

    251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
        |||||||||    ||||||||||||    |||    ||||||:::|||:::||
   1429 ACNACNCARYTNGARACNGAYAARYTNGCNYTNCCNATGWSNCCNWSNCC 1478

    267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
        |    :::|||:::|||||||||:::||||||||||||    ||||||:::|
   1479 NYTNWSNCCNWSNCCNATHCCNWSNCCNAAYGCNAARYTNGARAAYWSNG 1528

    284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
        ||        ||||||||||||:::||||||||||||||||||||||||||
   1529 CNYTNYTNACNGTNGARCCNWSNCCNCARGAYAARAAYAARGGNTTYTTY 1578

    301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
        |||||||||:::||||||            ||||||:::|||::::::|||::
   1579 GTNGAYGARWSNGARCCNYTNYTNMGNTGYGAYWSNACNWSNWSNGGNWS 1628

    317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
        ::::|||   :::    ||||||:::||||||||||||||||||||||||
```

```
    1629 NWSNGCNYTNWSNMGNAAYGGNWSNTTYATHACNAARGARAARAARGAYA 1678

     334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
         |||||      ||||||      ||||||||||||   ||||||||||||
    1679 CNGTNYTNMGNCARGTNMGNYTNGAYCCNTGYGAYYTNCARCCNATHTTY 1728

     351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
         |||||||||   ||||||||||||||||      ||||||||||
    1729 GAYGAYATGYTNCAYTTYYTNAAYCCNGARGARYTNMGNGTNATHGARGA 1778

     367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
         ||||||||||||||||||||   |||      ||||||||||||||||
    1779 RATHCCNCARGCNGARGAYAARYTNGAYMGNYTNTTYGARATHATHGGNG 1828

     384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
         |||||:::|||||||||:::|||||      |||:::|||||:::|||
    1829 TNAARWSNCARGARGCNWSNCARACNYTNYTNGAYWSNGTNTAYWSNCAY 1878

     401 LeuProAsp 403
         ||||||
    1879 YTNCCNGAY 1887

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX25322

seq_documentation_block:
ID   AAX25322 standard; cDNA; 2638 BP.
XX
AC   AAX25322;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Human TNF receptor TRL II cDNA.
XX
KW   TRL II; TNF receptor-like; tumour necrosis factor receptor; human;
KW   signal transduction; cell differentiation; prostate cancer;
KW   inflammation; arthritis; diabetes; insulin resistance; diagnosis;
KW   therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             510..2327
FT                   /*tag= a
FT   sig_peptide     510..632
FT                   /*tag= b
FT   mat_peptide     633..2324
XX
PN   WO9915663-A1.
XX
PD   01-APR-1999.
XX
PF   25-SEP-1998;   98WO-US20219.
XX
PR   17-MAR-1998;   98US-0042785.
PR   26-SEP-1997;   97US-0038896.
XX
PA   (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI   Busfield SJ;
XX
DR   WPI; 1999-254712/21.
DR   P-PSDB; AAY05697.
XX
PT   Nucleic acid encoding tumor necrosis factor receptor-like protein
XX
PS   Claim 1b; Page 148-153; 170pp; English.
XX
CC   This cDNA sequence codes for a novel member of the human tumour
CC   necrosis factor receptor superfamily (see AAY05697), termed TRL II
CC   (TNF receptor-like). A plasmid containing the full-length
CC   nucleotide sequence encoding human TRL II (clone ephT75L) is
CC   deposited as ATCC 98649). TRL II has a mol.wt. of 66.2 kDa.
CC   It is a membrane-bound protein that has 4 cysteine-rich domains
CC   and a C-terminal death domain (see AAY05698). TRL II cDNA was
CC   isolated from a HeLa cDNA library using mouse TRL cDNA (see
CC   AAX25320) as probe. An alternatively spliced, secreted form, human
CC   TRL I (see AAX25321), has also been identified. Chromosomal mapping
CC   located the TRL gene on chromosomal 6p25. The invention provides
CC   full-length TRL proteins, TRL fusion proteins, antigenic peptides
CC   and anti-TRL antibodies, as well as TRL nucleic acids, recombinant
CC   expression vectors, host cells and non-human transgenic animals.
CC   TRL are regulators or modulators of cellular signal transduction,
CC   cellular proliferation or differentiation, cell survival and
CC   apoptosis, immune system cells, and cells involved in insulin
CC   resistance or the diabetic response. They can be used to isolate
CC   cognate ligands, to modulate ligand/TRL interactions, to screen
CC   for potential modulators, and to treat conditions associated with
CC   abnormal TRL levels. TRL nucleic acids are used for recombinant
CC   production of TRL, to detect mutations, to identify TRL homologues,
CC   as antisense, triplex-forming or ribozyme therapeutics, for
CC   chromosome mapping and for forensic identification of individuals.
CC   Antibodies, antisense sequences, TRL protein mutants etc. that
CC   modulate activity of TRL are used to treat or prevent conditions
CC   associated with aberrant TRL protein or expression, e.g. cell
CC   proliferation and differentiation disorders (particularly cancer,
CC   especially of the prostate), inflammation (e.g. arthritis) and
CC   metabolic disease (e.g. diabetes and insulin resistance).
XX
SQ   Sequence 2638 BP; 625 A; 749 C; 701 G; 563 T; 0 other;


alignment_scores:
           Quality: 1527.00              Length:    294
             Ratio:  5.194                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AAX25322   ..

 Align seg 1/1  to: AAX25322  from: 1  to: 2638

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1260 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1309

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1310 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1359

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1360 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1409

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1410 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1459

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1460 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1509

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1510 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1559

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1560 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1609

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1610 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1659

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1660 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1709
```

```
 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1710 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1759

 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1760 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1809

 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1810 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1859

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1860 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1909

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1910 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1959

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1960 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 2009

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2010 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 2059

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2060 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 2109

 284 laLeuLeuThrValGluProSerProGlnAsp 294
     ||||||||||||||||||||||||||||||
2110 CTCTCCTGACGGTGGAGCCTTCCCCACAGGAT 2141
```

```
OM of: US-09-733-956-2 to: GenEmbl:* , out_format : pfs

Date: Jan 4, 2002  9:37 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09733956/runat_04012002_084025_15512/app_query.fasta_1.467
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09733956_@CGN1_1_5863 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-733-956-2
Query length: 405
Database: GenEmbl:*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 1599.630000


score_list:
Sequence              Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pr:AF068868          + 2091.00 3013.13  1.5e-159  1968  ! AF068868 Homo sapiens TNFR-rel
gb_pat:E29452           + 2091.00 3012.13  1.7e-159  2186  ! E29452 Tumor necrosis-associat
gb_pat:AX008646         + 2091.00 3010.24  2.1e-159  2666  ! AX008646 Sequence 1 from Paten
gb_pr:AK001504          + 2087.00 3004.55  4.4e-159  2636  ! AK001504 Homo sapiens cDNA FLJ
gb_pat:AR134329         + 1914.00 2758.11  2.4e-145  1719  ! AR134329 Sequence 5 from paten
gb_pat:AR134327         + 1914.00 2751.80  5.3e-145  3331  ! AR134327 Sequence 1 from paten
gb_ro:AY043489          + 1911.00 2752.48  4.9e-145  1968  ! AY043489 Mus musculus death re
gb_ro:AF322069          + 1898.00 2733.65  5.5e-144  1968  ! AF322069 Mus musculus DR6 mRNA
gb_ov:AF349908          + 1564.00 2246.37  7.6e-117  2879  ! AF349908 Gallus gallus death r
gb_pat:AR134340         + 1527.00 2197.19  4.2e-114  1815  ! AR134340 Sequence 24 from pate
gb_pat:AR134339         + 1527.00 2193.62  6.6e-114  2638  ! AR134339 Sequence 22 from pate
gb_pat:AR134328         + 1269.00 1820.12  4.2e-93   2612  ! AR134328 Sequence 3 from paten
gb_pr:HSJ181J13         +  874.00 1211.79  3.2e-59   118271  ! AL096801 Human DNA sequence
gb_htg:AL512350         -  874.00 1209.51  4.3e-59   150180  ! AL512350 Homo sapiens chromo
gb_htg:AC026260         -  408.00  536.43  1.3e-21   125419  ! AC026260 Homo sapiens chromo
gb_htg:AC026260         +  387.00  506.02  6.6e-20   125419  ! AC026260 Homo sapiens chromo
gb_pat:AX192854         +  339.00  493.91  3.1e-19   304  ! AX192854 Sequence 421 from Pate
gb_htg:AC069365         +  329.00  428.32  1.4e-15   64812  ! AC069365 Homo sapiens chromos
gb_ro:RNNGFRR           +  120.50  154.90  2.37      3259  ! X05137 Rat mRNA for fast nerve
gb_htg:AC019528         +  120.50  129.36  62.70     47525  ! AC019528 Drosophila melanogas
gb_in:AE003525          +  120.50  112.06  576.34    291829  ! AE003525 Drosophila melanoga
gb_pr:AB032251          +  119.00  142.33  11.88     9700  ! AB032251 Homo sapiens BPTF mRN
gb_in:AF216642          +  116.00  152.73  3.13      2065  ! AF216642 Drosophila melanogast
gb_ro:MMU52925          +  115.50  140.95  14.17     6585  ! U52925 Mus musculus coagulatic
gb_htg:AC015915         -  115.00  106.89  1.1e+03   217686  ! AC015915 Mus musculus chromo
gb_pr:HSM800344         +  114.50  155.80  2.11      1191  ! AL049996 Homo sapiens mRNA; cD
gb_pr:AK027184          +  114.00  150.25  4.30      1976  ! AK027184 Homo sapiens cDNA: FL
gb_ro:AF026504          +  113.50  138.33  19.84     6400  ! AF026504 Rattus norvegicus SPA
gb_ba:AE005150          -  113.50  133.15  38.54     11019  ! AE005150 Halobacterium sp. NR
gb_pat:AR153575         +  113.00  143.67  10.00     3386  ! AR153575 Sequence 1 from paten
gb_pr:HUMNGFR           +  113.00  143.67  10.00     3386  ! M14764 Human nerve growth fact
gb_vi:PPPVCG            +  112.50  132.83  40.14     9786  ! X81083 Plum pox potyvirus comp
gb_ba:PAU23764          +  112.00  152.11  3.39      1200  ! U23764 Pseudomonas aeruginosa
gb_pat:A43530           +  112.00  149.37  4.82      1600  ! A43530 Sequence 1 from Patent
gb_pat:AR098123         +  112.00  149.37  4.82      1600  ! AR098123 Sequence 1 from paten
gb_om:OCCALSEQ          +  112.00  145.08  8.35      2510  ! X55040 O.cuniculus mRNA for ca
gb_ba:AE004965          -  112.00  131.57  47.19     10353  ! AE004965 Pseudomonas aeruginc
gb_pr:AC010504          -  111.50  117.44  289.00    42265  ! AC010504 Homo sapiens chromos
gb_pr:AB007900          +  111.00  136.51  25.04     5296  ! AB007900 Homo sapiens KIAA044(
gb_pr:AF090989          +  111.00  135.38  28.96     5965  ! AF090989 Homo sapiens high-ris
gb_pr:AF090990          +  111.00  135.28  29.33     6028  ! AF090990 Homo sapiens high-ris
gb_ro:MMHC438N12        -  110.50  103.93  1.6e+03   149886  ! AF049850 Mus musculus maj
gb_ro:MMHC310M6         -  110.50  103.40  1.7e+03   158405  ! AF109906 Mus musculus MHC
gb_htg:AC008735         -  110.00   99.23  3.0e+03   227482  ! AC008735 Homo sapiens chr
gb_in:CEY75B8A          -  110.00   96.64  4.2e+03   298406  ! AL033514 Caenorhabditis e

seq_name: gb_pr:AF068868

seq_documentation_block:
LOCUS       AF068868     1968 bp    mRNA            PRI       05-SEP-1998
DEFINITION  Homo sapiens TNFR-related death receptor-6 (DR6) mRNA, complete
            cds.
ACCESSION   AF068868
VERSION     AF068868.1  GI:3549262
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1968)
  AUTHORS   Pan,G., Bauer,J.H., Haridas,V., Wang,S., Liu,D., Ni,J., Yu,G.,
            Vincenz,C., Aggarwal,B.B. and Dixit,V.M.
  TITLE     Identification and functional characterization of DR6, a novel
            death domain-containing TNF receptor
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1968)
  AUTHORS   Pan,G., Ni,J., Yu,G. and Dixit,V.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-1998) Pathology, University of Michgan Medical
            School, 1301 Catherine Road,, Ann Arbor, MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .1968
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .1968
                     /gene="DR6"
     CDS             1. .1968
                     /gene="DR6"
                     /codon_start=1
                     /product="TNFR-related death receptor-6"
                     /protein_id="AAC34583.1"
                     /db_xref="GI:3549263"
                     /translation="MGTSPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPE
                     QKASNLIGTYRHVDRATGQVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFTRHENG
                     IEKCHDCSQPCPWPMIEKLPCAALTDRECTCPPGMFQSNATCAPHTVCPVGWGVRKKG
                     TETEDVRCKQCARGTFSDVPSSVMKCKAYTDCLSQNLVVIKPGTKETDNVCGTLPSFS
                     SSTSPSPGTAIFPRPEHMETHEVPSSTYVPKGMNSTESNSSASVRPKVLSSIQEGTVP
                     DNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHILKLLPSMEATGGEKSSTPIKGPKRG
                     HPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRKSSRTLKKGPRQDPSAIVEKA
                     GLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSN
                     GYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDK
                     LALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLLRCDSTSSG
                     SSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQAEDK
                     LDRLFEIIGVKSQEASQTLLDSVYSHLPDLL"
BASE COUNT      479 a    583 c    512 g    394 t
ORIGIN


alignment_scores:
            Quality: 2091.00              Length:    405
              Ratio:  5.163                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AF068868   ..

 Align seg 1/1  to: AF068868  from: 1  to: 1968

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 800

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
 801 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 850

  34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 851 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 900

  51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 901 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 950

  67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 951 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1000

  84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1001 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1050

 101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1051 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1100

 117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1101 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1150

 134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1151 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1200

 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1201 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1250

 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1251 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1300

 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1301 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1350

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1351 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1400

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1401 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1450

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1451 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1500

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1501 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1550

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1551 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1600

 284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1601 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1650

 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1651 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1700

 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1701 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1750

 334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1751 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 1800

 351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1801 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 1850

 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1851 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 1900

 384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1901 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 1950

 401 LeuProAspLeuLeu 405
     ||||||||||||||||
1951 CTTCCTGACCTGCTG 1965
```

```
seq_name: gb_pat:E29452

seq_documentation_block:
LOCUS       E29452       2186 bp    DNA             PAT       07-FEB-2001
DEFINITION  Tumor necrosis-associated receptor TR7.
ACCESSION   E29452
VERSION     E29452.1  GI:13025564
KEYWORDS    JP 1999018786-A/1.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2186)
  AUTHORS   Keith,C.D.M.M., Hurle,P.Y.K.K. and Tan.
  TITLE     Tumor necrosis-associated receptor TR7
  JOURNAL   Patent: JP 1999018786-A 1 26-JAN-1999;
            SMITHKLINE BEECHAM CORP
COMMENT     OS   Unidentified
            PN   JP 1999018786-A/1
            PD   26-JAN-1999
            PF   02-APR-1998 JP 1998126559
            PR   02-APR-1997 US    60/041796,28-OCT-1997 US    08/959382 PI
            KEITH C DEAN,MARK R HURLE,PETER YOUNG,K B TAN PC
            C12N15/09,A61K38/00,A61K38/00,A61K38/00,A61K38/00,A61K45/00, PC
            A61K48/00,
            PC   C07K14/705,C07K16/30,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC
            C12N5/10,
            PC   C12P21/02,G01N33/53,G01N33/531,G01N33/566,G01N33/577,
            C12N15/00, PC   A61K37/02,
            PC   A61K37/02,A61K37/02,A61K37/02,C12N5/00,C12N5/00 CC
            Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .2186
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2186
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      514 a    613 c    565 g    494 t
ORIGIN


alignment_scores:
           Quality: 2091.00            Length:    405
             Ratio:  5.163               Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x E29452   ..

 Align seg 1/1  to: E29452  from: 1  to: 2186
```

```
   1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 757 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 806

  17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 807 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 856

  34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 857 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 906

  51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 907 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 956

  67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 957 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1006

  84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1007 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1056

 101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1057 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1106

 117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1107 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1156

 134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1157 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1206

 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1207 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1256

 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1257 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1306

 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1307 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1356

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1357 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1406

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1407 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1456

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1457 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1506

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1507 ACCACCCAGCTGGAAACTGACAAACTAGCTCTGCCGATGAGCCCCAGCCC 1556

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1557 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1606

 284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1607 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1656

 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1657 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1706

 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1707 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1756

 334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1757 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 1806

 351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1807 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 1856

 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1857 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 1906

 384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1907 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 1956

 401 LeuProAspLeuLeu 405
     ||||||||||||||||
1957 CTTCCTGACCTGCTG 1971
```

```
seq_name: gb_pat:AX008646

seq_documentation_block:
LOCUS       AX008646     2666 bp    DNA             PAT       06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9966039.
ACCESSION   AX008646
VERSION     AX008646.1  GI:9996171
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2666)
  AUTHORS   Fieles,W.E., Gomes,B.C. and Lu,J.J.
  TITLE     Tnfr-like protein with death domain
  JOURNAL   Patent: WO 9966039-A 1 23-DEC-1999;
            ZENECA LTD (GB); FIELES WILLIAM EDWARD (US); GOMES BRUCE CHARLES
            (US); LU JIAN JASON (US)
FEATURES             Location/Qualifiers
     source          1. .2666
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      691 a    609 c    591 g    775 t
ORIGIN


alignment_scores:
            Quality: 2091.00             Length:    405
              Ratio:  5.163                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AX008646   ..

 Align seg 1/1  to: AX008646  from: 1  to: 2666
```

```
   1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 190 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 239

  17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 240 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 289

  34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
 290 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 339

  51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl  67
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 340 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 389

  67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG  84
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 390 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 439

  84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 440 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 489

 101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 490 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 539

 117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 540 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 589

 134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 590 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 639

 151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 640 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 689

 167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 690 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 739

 184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 740 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 789

 201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 790 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 839

 217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 840 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 889

 234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 890 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 939

 251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 940 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 989

 267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 990 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1039

 284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1040 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 1089

 301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1090 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 1139

 317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1140 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 1189

 334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1190 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 1239

 351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1240 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 1289

 367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1290 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 1339

 384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1340 TCAAGAGCCAGGAAGCCAGCCAGACCCTCCTGGACTCTGTTTATAGCCAT 1389

 401 LeuProAspLeuLeu 405
     |||||||||||||||
1390 CTTCCTGACCTGCTG 1404
```

seq_name: gb_pr:AK001504

seq_documentation_block:

```
LOCUS       AK001504     2636 bp    mRNA            PRI       22-FEB-2000
DEFINITION  Homo sapiens cDNA FLJ10642 fis, clone NT2RP2005752, highly similar
            to Homo sapiens TNFR-related death receptor-6 mRNA.
ACCESSION   AK001504
VERSION     AK001504.1  GI:7022799
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens teratocarcinoma cell_line:NT2 cDNA to mRNA,
            clone_lib:NT2RP2 clone:NT2RP2005752.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K.,
            Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Ishii,S., Kawai,Y.,
            Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K.,
            Masuho,Y. and Kanehori,K.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished (2000)
REFERENCE   2  (bases 1 to 2636)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao
            Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,
            Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .2636
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NT2"
                     /cell_type="teratocarcinoma"
                     /clone="NT2RP2005752"
                     /clone_lib="NT2RP2"
                     /note="cloning vector: pME18SFL3; mRNA from NT2 neuronal
                     precursor cells after 2-weeks retinoic acid (RA)
                     induction."
BASE COUNT      569 a    762 c    722 g    583 t
ORIGIN
```

alignment_scores:

```
            Quality: 2087.00              Length:    405
              Ratio:  5.153                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 99.753
```

```
alignment_block:
 US-09-733-956-2 x AK001504   ..

 Align seg 1/1  to: AK001504  from: 1  to: 2636

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1145 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1194

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1195 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1244

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1245 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1294

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1295 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1344

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1345 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1394

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1395 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1444

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1445 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1494

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1495 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1544

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1545 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1594

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1595 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1644

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1645 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1694

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1695 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1744

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1745 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1794

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1795 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1844

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1845 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1894

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1895 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1944

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1945 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1994

     284 laLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhePhe 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1995 CTCTCCTGACGGTGGAGCCTTCCCCACAGGACAAGAACAAGGGCTTCTTC 2044

     301 ValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGlySe 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2045 GTGGATGAGTCGGAGCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTC 2094

     317 rSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysAspT 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2095 CTCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACA 2144

     334 hrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIlePhe 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2145 CAGTGTTGCGGCAGGTACGCCTGGACCCCTGTGACTTGCAGCCTATCTTT 2194

     351 AspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGluGl 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2195 GATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGA 2244

     367 uIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleGlyV 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    2245 GATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAG 2294

     384 alLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSerHis 400
         ||||||||||||||||||||||||:::||||||||||||||||||||||||
    2295 TCAAGAGCCAGGAAGCCAGCCAGACCTTCCTGGACTCTGTTTATAGCCAT 2344

     401 LeuProAspLeuLeu 405
         |||||||||||||||
    2345 CTTCCTGACCTGCTG 2359

seq_name: gb_pat:AR134329

seq_documentation_block:
LOCUS       AR134329     1719 bp    DNA             . PAT       16-MAY-2001
DEFINITION  Sequence 5 from patent US 6194151.
ACCESSION   AR134329
VERSION     AR134329.1  GI:14123234
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1719)
  AUTHORS   Busfield,S.J.
  TITLE     Molecules of the TNF receptor superfamily and uses therefor
  JOURNAL   Patent: US 6194151-A 5 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .1719
                     /organism="unknown"
BASE COUNT      419 a    503 c    457 g    340 t
ORIGIN


alignment_scores:
            Quality: 1914.00            Length:    406
              Ratio:  4.883              Gaps:      2
 Percent Similarity:  96.552   Percent Identity: 92.857

alignment_block:
 US-09-733-956-2 x AR134329   ..

 Align seg 1/1  to: AR134329  from: 1  to: 1719

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||:::||||||||||:::||||||||||||   ||||||
     505 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 554

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
          |||:::|||:::|||||||||||||||||||||||||||||||:::   |
     555 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 604

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
```

```
          |||||||    :::|||:::|||||||||||          |||:::|||
      605 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 654

       51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
          ||||||:::||||||||||||||||||||||||||||||||||    |||||
      655 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 704

       66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
          |||||||||   |||||||||||||||||   ||||||:::||||||
      705 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 751

       83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
          |||||||||||||||||||  |||||||||||||||||||||||||||
      752 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 801

      100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
          ||||||||||||||||||||||||||||||||||||||||||||:::|||||
      802 TTGCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGT 851

      116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      852 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 901

      133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
          ||||||||||||||||||||||||||||||||||||||||||||:::|||
      902 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 951

      150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
          ||||||||||||||||||||||||||||||||||  |||||||||||||
      952 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 1001

      166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1002 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACA 1051

      183 leTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1052 TCTATCAGTTTCTTTGCAACGCCAGTGAGAGGGAGGTGGCGGCCTTCTCC 1101

      200 AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1102 AATGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCAGCACTG 1151

      216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1152 GACCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGC 1201

      233 rgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1202 GCCAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAA 1251

      250 AspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSe 266
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1252 GACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAGCCCCAG 1301

      266 rProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnS 283
          ||||||||||||||||||||||||||||||||||||:::|||||||||||||
     1302 TCCGCTGAGCCCGAGCCCCATCCCCAGTCCTAACGTGAAACTTGAGAATT 1351

      283 erAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
          ||:::||||||||||||||||||||||||   ||||||||||||   |||
     1352 CCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAGTGCTTC 1401

      300 PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGl 316
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1402 TTCGTGGACGAGTCAGAGCCCCTTCTGCGTTGCGACTCCACATCCAGTGG 1451

      316 ySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysA 333
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1452 CTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAAGAAAAGAAGG 1501

      333 spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1502 ACACAGTGTTGCGGCAGGTCCGCCTGGACCCCTGTGACTTGCAGCCCATC 1551

      350 PheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGl 366
          ||||||||||||||||||:::|||||||||||||||||||||||||||||
     1552 TTTGATGACATGCTGCATATCCTGAACCCCGAGGAGCTGCGGGTGATTGA 1601

      366 uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleG 383
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1602 AGAGATTCCCCAGGCTGAGGACAAACTGGACCGCCTCTTCGAGATCATTG 1651

      383 lyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSer 399
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1652 GGGTCAAGAGCCAAGAAGCCAGCCAGACCCTCTTGGACTCTGTGTACAGT 1701

      400 HisLeuProAspLeuLeu 405
          ||||||||||||||||||
     1702 CATCTTCCTGACCTATTG 1719

seq_name: gb_pat:AR134327

seq_documentation_block:
LOCUS       AR134327      3331 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 1 from patent US 6194151.
ACCESSION   AR134327
VERSION     AR134327.1  GI:14123232
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3331)
  AUTHORS   Busfield,S.J.
  TITLE     Molecules of the TNF receptor superfamily and uses therefor
  JOURNAL   Patent: US 6194151-A 1 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .3331
                     /organism="unknown"
BASE COUNT      825 a    891 c    739 g    876 t
ORIGIN


alignment_scores:
            Quality: 1914.00             Length:     406
              Ratio:   4.883               Gaps:       2
 Percent Similarity:  96.552   Percent Identity: 92.857

alignment_block:
 US-09-733-956-2 x AR134327   ..

 Align seg 1/1  to: AR134327  from: 1  to: 3331

        1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
          ||||||||||||:::|||||||||:::|||||||||||||    ||||||
      848 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 897

       17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
           |||:::|||:::||||||||||||||||||||||||||||||:::    |
      898 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 947

       34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
          ||||||||    :::|||:::|||||||||||          |||:::|||
      948 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 997

       51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
          ||||||:::||||||||||||||||||||||||||||||||||    |||||
      998 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 1047

       66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
          |||||||||   |||||||||||||||||   ||||||:::||||||
     1048 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 1094

       83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
          |||||||||||||||||||   ||||||||||||||||||||||||||||
     1095 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 1144
```

```
 100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
     ||||||||||||||||||||||||||||||||||||||||:::|||||
1145 TTGCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGT 1194

 116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1195 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 1244

 133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
     ||||||||||||||||||||||||||||||||||||||||:::|||
1245 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 1294

 150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
     |||||||||||||||||||||||||||||||  ||||||||||||||
1295 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 1344

 166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1345 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACA 1394

 183 leTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1395 TCTATCAGTTTCTTTGCAACGCCAGTGAGAGGGAGGTGGCGGCCTTCTCC 1444

 200 AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1445 AATGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCAGCACTG 1494

 216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1495 GACCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGC 1544

 233 rgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1545 GCCAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAA 1594

 250 AspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSe 266
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1595 GACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAGCCCCAG 1644

 266 rProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnS 283
     ||||||||||||||||||||||||||||||||:::||||||||||||||
1645 TCCGCTGAGCCCGAGCCCCATCCCCAGTCCTAACGTGAAACTTGAGAATT 1694

 283 erAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
     ||:::|||||||||||||||||||||||||    ||||||||||||   |||
1695 CCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAGTGCTTC 1744

 300 PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGl 316
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1745 TTCGTGGACGAGTCAGAGCCCCTTCTGCGTTGCGACTCCACATCCAGTGG 1794

 316 ySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysA 333
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1795 CTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAAGAAAAGAAGG 1844

 333 spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1845 ACACAGTGTTGCGGCAGGTCCGCCTGGACCCCTGTGACTTGCAGCCCATC 1894

 350 PheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGl 366
     ||||||||||||||||||:::||||||||||||||||||||||||||||
1895 TTTGATGACATGCTGCATATCCTGAACCCCGAGGAGCTGCGGGTGATTGA 1944

 366 uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleG 383
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1945 AGAGATTCCCCAGGCTGAGGACAAACTGGACCGCCTCTTCGAGATCATTG 1994

 383 lyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSer 399
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1995 GGGTCAAGAGCCAAGAAGCCAGCCAGACCCTCTTGGACTCTGTGTACAGT 2044

     400 HisLeuProAspLeuLeu 405
         ||||||||||||||||||
    2045 CATCTTCCTGACCTATTG 2062

seq_name: gb_ro:AY043489

seq_documentation_block:
LOCUS       AY043489     1968 bp    mRNA            ROD       25-JUL-2001
DEFINITION  Mus musculus death receptor 6 mRNA, complete cds.
ACCESSION   AY043489
VERSION     AY043489.1  GI:15020325
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1968)
  AUTHORS   Kim,V., Machleidt,T., Shi,W.-X., Wang,X. and Cai,Z.
  TITLE     Murine DR6: murine TNFR-related death receptor-6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1968)
  AUTHORS   Kim,V., Machleidt,T., Shi,W.-X., Wang,X. and Cai,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Immunology, RW Johnson Pharmaceutical
            Research Institute, 3210 Merryfield Row, San Diego, CA 92121, USA
FEATURES             Location/Qualifiers
     source          1. .1968
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /chromosome="17"
                     /sex="male"
                     /tissue_type="kidney"
                     /dev_stage="9-11 weeks"
     CDS             1. .1968
                     /note="TNFR-related DR6"
                     /codon_start=1
                     /product="death receptor 6"
                     /protein_id="AAK74193.1"
                     /db_xref="GI:15020326"
                     /translation="MGTRASSITALASCSRTAGQVGATMVAGSLLLLGFLSTITAQPE
                     QKTLSLPGTYRHVDRTTGQVLTCDKCPAGTYVSEHCTNMSLRVCSSCPAGTFTRHENG
                     IERCHDCSQPCPWPMIERLPCAALTDRECICPPGMYQSNGTCAPHTVCPVGWGVRKKG
                     TENEDVRCKQCARGTFSDVPSSVMKCKAHTDCLGQNLEVVKPGTKETDNVCGMRLFFS
                     STNPPSSGTVTFSHPEHMESHDVPSSTYEPQGMNSTDSNSTASVRTKVPSGIEEGTVP
                     DNTSSTSGKEGTNRTLPNPPQVTHQQAPHHRHILKLLPSSMEATGEKSSTAIKAPKRG
                     HPRQNAHKHFDINEHLPWMIVLFLLLVLVLIVVCSIRKSSRTLKKGPRQDPSAIVEKA
                     GLKKSLTPTQNREKWIYYRNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSN
                     GYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDK
                     LALPMSPSPLSPSPMPSPNVKLENSTLLTVEPSPLDKNKCFFVDESEPLLRCDSTSSG
                     SSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHILNPEELRVIEEIPQAEDK
                     LDRLFEIIGVKSQEASQTLLDSVYSHLPDLL"
BASE COUNT      474 a    594 c    517 g    383 t
ORIGIN


alignment_scores:
            Quality: 1911.00             Length:    406
              Ratio:  4.875                Gaps:      2
 Percent Similarity:  96.552   Percent Identity: 92.611

alignment_block:
 US-09-733-956-2 x AY043489   ..

 Align seg 1/1  to: AY043489  from: 1  to: 1968

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||:::||||||||||:::|||||||||||   ||||||
     751 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 800

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
           |||:::|||:::|||||||||||||||||||||||||||||:::   |
     801 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 850
```

```
 34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
    ||||||||   :::|||:::|||||||||||||         |||:::|||
851 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACAAGTTACCCAC 900

 51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
    ||||||:::|||||||||||||||||||||||||||||||   |||||
901 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 950

 66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
    |||||||||   |||||||||||||||||   ||||||:::|||||||
951 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 997

 83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
    ||||||||||||||||||   |||||||||||||||||||||||||||||
998 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 1047

100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
    ||||||||||||||||||||||||||||||||||||||||||:::|||||
1048 TTGCCTTGGATGATCGTCCTCTTCCTTCTGCTGGTCCTGGTGCTGATAGT 1097

116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
    |||||||||||||||||||||||||||||||||||||||||||||||||
1098 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 1147

133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
    |||||||||||||||||||||||||||||||||||||||||||:::|||
1148 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 1197

150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
    |||||||||||||||||||||||||||||||   ||||||||||||||
1198 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 1247

166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
    |||||||||||||||||||||||||||||||||||||||||||||||||
1248 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACA 1297

183 leTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
    |||||||||||||||||||||||||||||||||||||||||||||||||
1298 TCTATCAGTTTCTTTGCAACGCCAGCGAGAGGGAGGTGGCGGCCTTCTCC 1347

200 AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
    |||||||||||||||||||||||||||||||||||||||||||||||||
1348 AATGGATACACTGCAGATCATGAACGGGGCTACGCGGCTCTGCAGCACTG 1397

216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
    |||||||||||||||||||||||||||||||||||||||||||||||||
1398 GACCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGC 1447

233 rgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
    |||||||||||||||||||||||||||||||||||||||||||||||||
1448 GCCAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAA 1497

250 AspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSe 266
    |||||||||||||||||||||||||||||||||||||||||||||||||
1498 GACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAGCCCCAG 1547

266 rProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnS 283
    |||||||||||||||||:::|||||||||||||:::|||||||||||||
1548 TCCGCTTAGCCCGAGCCCCATGCCCAGTCCTAACGTGAAACTTGAGAATT 1597

283 erAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
    ||:::|||||||||||||||||||||   ||||||||||||   |||
1598 CCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAGTGCTTC 1647

300 PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGl 316
    |||||||||||||||||||||||||||||||||||||||||||||||||
1648 TTCGTGGACGAGTCAGAGCCCCTTCTGCGATGCGACTCCACATCCAGTGG 1697

316 ySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysA 333
    |||||||||||||||||||||||||||||||||||||||||||||||||
1698 CTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAAGAAAAGAAGG 1747

333 spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
    |||||||||||||||||||||||||||||||||||||||||||||||||
1748 ACACAGTGTTGCGGCAGGTCCGCCTGGACCCCTGTGACTTGCAGCCCATC 1797

350 PheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGl 366
    ||||||||||||||||||:::|||||||||||||||||||||||||||||
1798 TTTGATGACATGCTGCATATCCTGAACCCCGAGGAGCTGCGGGTGATTGA 1847

366 uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleG 383
    |||||||||||||||||||||||||||||||||||||||||||||||||
1848 AGAGATTCCCCAGGCTGAGGACAAACTGGACCGCCTCTTCGAGATCATTG 1897

383 lyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSer 399
    |||||||||||||||||||||||||||||||||||||||||||||||||
1898 GGGTCAAGAGCCAAGAAGCCAGCCAGACCCTCTTGGACTCTGTGTACAGT 1947

400 HisLeuProAspLeuLeu 405
    ||||||||||||||||||
1948 CATCTTCCTGACCTATTG 1965
```

seq_name: gb_ro:AF322069

seq_documentation_block:

```
LOCUS       AF322069     1968 bp    mRNA            ROD       06-DEC-2000
DEFINITION  Mus musculus DR6 mRNA, complete cds.
ACCESSION   AF322069
VERSION     AF322069.1  GI:11559849
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1968)
  AUTHORS   Isogai,D., Ichino,M., Yoshinari,M., Yamaura,A., Kurokawa,F. and
            Minami,M.
  TITLE     Mouse DR6: mouse homolog of human TNFR-related death receptor-6
            (DR6)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1968)
  AUTHORS   Isogai,D., Ichino,M., Yoshinari,M., Yamaura,A., Kurokawa,F. and
            Minami,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2000) Immunology, Yokohama City University School
            of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236-0004,
            Japan
FEATURES             Location/Qualifiers
     source          1. .1968
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /tissue_type="kidney"
     CDS             1. .1968
                     /codon_start=1
                     /product="DR6"
                     /protein_id="AAG38115.1"
                     /db_xref="GI:11559850"
                     /translation="MGTRASSITALASCSRTAGQVGATMVAGSLLLLGFLSTITAQPE
                     QKTLSLPGTYRHVDRTTGQVLTCDKCPAGTYVSEHCTNMSLRVCSSCPAGTFTRHENG
                     IERCHDCSQPCPWPMIERLPCAALTDRECICPPGMYQSNGTCAPHTVCPVGWGVRKKG
                     TENEDVRCKQCARGTFSDVPSSVMKCKAHTDCLGQNLEVVKPGTKETDNVCGMRLFFS
                     STNPPSSGTVTFSHPEHMESHDVPSSTYEPQGMNSTDSNSTASVRTKVPSGIEEGTVP
                     DNTSSTSGKEGTNRTLPNPPQVTHQQAPHHRHILKLLPSSMEATGEKSSTAIKAPKRG
                     HPRQNAHKHFDINEHLPLMIVLFLLLVLVLIVVCSIRKSSRTLKKGPRQDPSAIVEKA
                     GLKKSLTPTQNREKWIYYRNGHGIDILKLVAAQVGSQWKDIYQFLCNASEREVAAFSN
                     GYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDTTQLETDK
                     LALPMSPSPLSPSPMPSPNVKLENSTLLTVEPSPLDKNKCFFVDESEPLLRCDSTSSG
                     SSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHILNPEELRVIEEIPQAEDK
                     LDRLFEIIGVKSQEASQTLLDSVYSHLPDLL"
BASE COUNT      475 a    594 c    514 g    385 t
ORIGIN
```

alignment_scores:

```
          Quality: 1898.00              Length:    406
            Ratio:  4.854                 Gaps:      2
Percent Similarity:  96.305   Percent Identity: 92.365

alignment_block:
US-09-733-956-2 x AF322069   ..

Align seg 1/1  to: AF322069  from: 1  to: 1968

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||:::||||||||||:::|||||||||||   ||||||
     751 ATGAACTCAACAGATTCCAACTCTACTGCCTCTGTTAGAACTAAGGTACC 800

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         |||:::|||:::|||||||||||||||||||||||||||||:::   |
     801 AAGTGGCATCGAGGAAGGGACAGTGCCTGACAATACGAGCTCAACCAGTG 850

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||   :::|||:::||||||||||||||         |||:::|||
     851 GGAAGGAAGGCACTAATAGGACCCTGCCAAACCCACCACACAAGTTACCCAC 900

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuPro...SerMe 66
         ||||||:::|||||||||||||||||||||||||||||||||   |||||
     901 CAGCAAGCCCCCCACCACAGACACATTCTGAAGCTGCTGCCATCGTCCAT 950

      66 tGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysA 83
         |||||||||   ||||||||||||||||||   ||||||:::|||||||
     951 GGAGGCCACG...GGTGAGAAGTCCAGCACAGCCATCAAGGCCCCCAAGA 997

      83 rgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHis 99
         ||||||||||||||||||   ||||||||||||||||||||||||||||
     998 GGGGTCACCCCAGACAGAACGCTCACAAGCATTTCGACATCAACGAGCAC 1047

     100 LeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleVa 116
         ||||||   ||||||||||||||||||||||||||||||||||:::||||||
    1048 TTGCCCTTGATGATCGTTCTCTTTCTTCTGCTTGTCCTGGTGCTGATAGT 1097

     116 lValCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgG 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1098 GGTGTGCAGTATCCGAAAGAGCTCCAGGACTCTCAAAAAGGGGCCCCGGC 1147

     133 lnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThr 149
         |||||||||||||||||||||||||||||||||||||||||||:::|||
    1148 AGGATCCCAGCGCCATAGTGGAAAAGGCGGGGCTGAAGAAGTCCCTGACT 1197

     150 ProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIl 166
         ||||||||||||||||||||||||||||||||   ||||||||||||||
    1198 CCCACCCAGAACCGGGAGAAATGGATCTACTACCGCAACGGCCATGGTAT 1247

     166 eAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspI 183
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1248 TGACATCTTGAAGCTTGTAGCAGCCCAGGTGGGAAGCCAGTGGAAGGACA 1297

     183 leTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSer 199
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1298 TCTATCAGTTTCTTTGCAACGCCAGCGAGAGGGAGGTGGCGGCCTTCTCC 1347

     200 AsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTr 216
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1348 AATGGATACACTGCAGATCATGAACGGGCCTACGCGGCTCTGCAGCACTG 1397

     216 pThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuA 233
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1398 GACCATCCGTGGCCCTGAGGCCAGCCTTGCCCAGCTCATTAGCGCCTTGC 1447

     233 rgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGlu 249
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1448 GCCAGCACCGACGCAATGATGTTGTGGAGAAGATTCGTGGGCTGATGGAA 1497

     250 AspThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSe 266
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1498 GACACCACGCAGTTGGAAACAGACAAACTGGCTCTCCCCATGAGCCCCAG 1547

     266 rProLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnS 283
         ||||||||||||||||||||:::||||||||||||:::||||||||||||
    1548 TCCGCTTAGCCCGAGCCCCATGCCCAGTCCTAACGTGAAACTTGAGAATT 1597

     283 erAlaLeuLeuThrValGluProSerProGlnAspLysAsnLysGlyPhe 299
         ||:::|||||||||||||||||||||||   ||||||||||||   |||
    1598 CCACTCTCCTGACAGTGGAGCCCTCACCGCTGGACAAGAACAAGTGCTTC 1647

     300 PheValAspGluSerGluProLeuLeuArgCysAspSerThrSerSerGl 316
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1648 TTCGTGGACGAGTCAGAGCCCCTTCTGCGATGCGACTCCACATCCAGTGG 1697

     316 ySerSerAlaLeuSerArgAsnGlySerPheIleThrLysGluLysLysA 333
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1698 CTCTTCAGCACTGAGCAGAAACGGCTCCTTTATTACCAAAGAAAAGAAGG 1747

     333 spThrValLeuArgGlnValArgLeuAspProCysAspLeuGlnProIle 349
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1748 ACACAGTGTTGCGGCAGGTCCGCCTGGACCCCTGTGACTTGCAGCCCATC 1797

     350 PheAspAspMetLeuHisPheLeuAsnProGluGluLeuArgValIleGl 366
         ||||||||||||||||||:::||||||||||||||||||||||||||||
    1798 TTTGATGACATGCTGCATATCCTGAACCCCGAGGAGCTGCGGGTGATTGA 1847

     366 uGluIleProGlnAlaGluAspLysLeuAspArgLeuPheGluIleIleG 383
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1848 AGAGATTCCCCAGGCTGAGGACAAACTGGACCGCCTCTTCGAGATCATTG 1897

     383 lyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSerValTyrSer 399
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1898 GGGTCAAGAGCCAAGAAGCCAGCCAGACCCTCTTGGACTCTGTGTACAGT 1947

     400 HisLeuProAspLeuLeu 405
         ||||||||||||||||||
    1948 CATCTTCCTGACCTATTG 1965

seq_name: gb_ov:AF349908

seq_documentation_block:
LOCUS       AF349908     2879 bp    mRNA            VRT       15-MAY-2001
DEFINITION  Gallus gallus death receptor 6 mRNA, complete cds.
ACCESSION   AF349908
VERSION     AF349908.2  GI:14043014
KEYWORDS    .
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 2879)
  AUTHORS   Bridgham,J.T. and Johnson,A.L.
  TITLE     Expression of DR6 in the ovary
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2879)
  AUTHORS   Bridgham,J.T. and Johnson,A.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2001) Biological Sciences, University of Notre
            Dame, P.O. Box 369, Notre Dame, IN 46556, USA
REFERENCE   3  (bases 1 to 2879)
  AUTHORS   Bridgham,J.T. and Johnson,A.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2001) Biological Sciences, University of Notre
            Dame, P.O. Box 369, Notre Dame, IN 46556, USA
  REMARK    Sequence update by submitter
COMMENT     On May 15, 2001 this sequence version replaced gi:13549145.
FEATURES             Location/Qualifiers
     source          1. .2879
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
     CDS             1. .1956
                     /note="DR6; death domain-containing tumor necrosis
                     receptor family member"
                     /codon_start=1
```

```
                     /product="death receptor 6"
                     /protein_id="AAK29666.2"
                     /db_xref="GI:14043015"
                     /translation="MGAAVLAAVLPLLVFLGTADAQPKLTSEQNAVSLPAGKYLHLDR
                     ATNQELICDKCPAGTYVSKHCTKSTLRECSPCPDGTFTKHENGIERCHPCRKPCELPM
                     IEKTHCTALTDRECTCLSGTFQINDTCVPYTVCPVGWGVRKKGTETEDVRCKPCLRGT
                     FSDVPSSVMKCKTYTDCFGKNMVVVKPGTKESDNVCGSPASLPNTSLTSSDAQADGET
                     YEAPPTAYLPKGLNSSVFDLSSSPAPRVSNGTAEPTVDYNDTSANGTVGAPGSLSSAG
                     TAGQAQSYRHKHTSQAMGKQPAQEMAGGEKSSIPYRPPRRGPPNVHQHFDINEHLPWM
                     IVLFLLLVLVVIVVCSVRKSSRTLKKGPRQDPSAIVEKAIMKKSTTPTQNREKWIYYC
                     NGHGIDILKPVAAQVGSQWKDIYQFLCNASEREVAAFSNGYAADHERAYAALQHWTIR
                     GPEASLAQLISALRQHRRNDVVEKIRGLMEDTTPVQMQPQWQTQDCSNDDGKLEGDKL
                     ALPVSPSPLSPVPTPSPKPPDAAVLTVEPSPSEKKCFFVDESEPLLRCDSTSSGSSAL
                     SRTGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHILNPEELHVIEEIPQAEDKLDRL
                     FEIAGVKSQEASQTLLDSVYSHLPDLL"
BASE COUNT      661 a    853 c    762 g    598 t      5 others
ORIGIN


alignment_scores:
            Quality: 1564.00              Length:    426
              Ratio:  4.344                Gaps:      5
 Percent Similarity:  84.507    Percent Identity: 74.883

alignment_block:
 US-09-733-956-2 x AF349908    ..

 Align seg 1/1  to: AF349908  from: 1  to: 2879

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         :::|||||||:::      :::   |||:::|||      |||:::|||
     691 CTGAACTCTTCCGTTTTTGATCTCTCCTCCTCTCCAGCGCCGCGTGTGTC 740

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ::::::      |||  ||||||      |||   :::||||||:::|
     741 CAATGGCACAGCAGAGCCAACAGTCGACTACAATGACACCTCGGCCAACG 790

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||       :::      :::|||   :::      ::::::::::::
     791 GCACAGTTGGTGCTCCTGGGAGCCTCTCCAGCGCCGGCACCGCTGGCCAG 840

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeu.........Pr 64
         |||:::   :::|||:::|||      :::   :::          ||
     841 GCACAGAGCTACCGGCACAAGCACACCAGCCAAGCAATGGGCAAGCAGCC 890

      64 oSerMetGluAlaThrGlyGlyGluLysSerSerThrProIleLysGlyP 81
         |:::::::|||   :::|||||||||||||||||||   |||   :::   |
     891 GGCGCAGGAGATGGCCGGGGGTGAGAAGTCCAGCATCCCCTACAGGCCGC 940

      81 roLysArgGlyHisProArgGlnAsnLeuHisLysHisPheAspIleAsn 97
         ||:::|||||   |||      |||:::|||:::|||||||||||||||
     941 CCCGGAGAGGGCCGCCC......AACGTCCACCAGCACTTCGACATCAAT 984

      98 GluHisLeuProTrpMetIleValLeuPheLeuLeuLeuValLeuValVa 114
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     985 GAACACTTGCCCTGGATGATCGTCCTCTTCTTGCTGCTGGTCCTGGTGGT 1034

     114 lIleValValCysSerIleArgLysSerSerArgThrLeuLysLysGlyP 131
         ||||||||||||||||:::||||||||||||||||||||||||||||||
    1035 CATCGTTGTCTGCAGCGTCAGGAAGAGCTCACGGACTCTGAAGAAGGGAC 1084

     131 roArgGlnAspProSerAlaIleValGluLysAlaGlyLeuLysLysSer 147
         |||||||||||||||||||||||||||||||   :::|||||||||
    1085 CGAGGCAAGATCCCAGTGCCATCGTGGAGAAAGCTATTATGAAAAAATCC 1134

     148 MetThrProThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHi 164
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1135 ACCACCCCCACACAGAACAGGGAGAAATGGATCTACTACTGCAACGGGCA 1184

     164 sGlyIleAspIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpL 181
         ||||||||||||||||||   ||||||||||||||||||||||||||||
    1185 TGGCATCGACATCCTGAAGCCTGTGGCAGCCCAGGTGGGGAGCCAGTGGA 1234

     181 ysAspIleTyrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAla 197
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1235 AGGACATCTACCAGTTCCTGTGCAATGCCAGTGAGCGCGAGGTGGCCGCC 1284

     198 PheSerAsnGlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGl 214
         ||||||||||||||:::||||||||||||||||||||||||||||||||
    1285 TTCTCCAACGGCTACGCGGCCGACCACGAGCGGGCCTACGCCGCGCTGCA 1334

     214 nHisTrpThrIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerA 231
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1335 GCACTGGACCATCCGCGGGCCTGAGGCCAGCCTGGCACAGCTCATCAGCG 1384

     231 laLeuArgGlnHisArgArgAsnAspValValGluLysIleArgGlyLeu 247
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1385 CGCTCCGCCAGCACCGCCGCAACGACGTGGTGGAAAAGATCCGCGGCCTG 1434

     248 MetGluAspThrThr.................................. 252
         ||||||||||||||||
    1435 ATGGAGGACACCACGCCGGTGCAGATGCAGCCTCAGTGGCAAACACAGGA 1484

     253 ..................GlnLeuGluThrAspLysLeuAlaLeuProM 263
                            :::||||||   ||||||||||||||||||:
    1485 CTGCTCAAATGATGACGGAAAGCTGGAAGGTGACAAACTGGCACTGCCTG 1534

     263 etSerProSerProLeuSerProSerProIleProSerProAsnAlaLys 279
         ::||||||||||||||||||||   |||   |||||||||      |||
    1535 TGAGCCCCAGCCCGCTCAGCCCGGTGCCCACCCCCAGCCCC......AAA 1578

     280 LeuGluAsnSerAlaLeuLeuThrValGluProSerProGlnAspLysAs 296
               ::::::|||:::||||||||||||||||||||::::::|||
    1579 CCACCCGATGCGGCCGTCCTCACCGTGGAGCCCTCTCCTTCAGAGAAG.. 1626

     296 nLysGlyPhePheValAspGluSerGluProLeuLeuArgCysAspSerT 313
          |||   ||||||||||||||||||||||||||||||||||||||||||
    1627 .AAATGCTTCTTTGTGGACGAGTCGGAGCCCCTCCTGCGGTGTGACTCCA 1675

     313 hrSerSerGlySerSerAlaLeuSerArgAsnGlySerPheIleThrLys 329
         ||||||||||||||||||||||||||||:::||||||||||||||||||
    1676 CCTCCAGCGGCTCCTCTGCGCTCAGCCGGACCGGCTCCTTTATTACCAAA 1725

     330 GluLysLysAspThrValLeuArgGlnValArgLeuAspProCysAspLe 346
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1726 GAAAAGAAGGACACAGTCTTGCGCCAGGTGCGCTTGGACCCTTGTGACCT 1775

     346 uGlnProIlePheAspAspMetLeuHisPheLeuAsnProGluGluLeuA 363
         ||||||||||||||||||||||||||:::||||||||||||||||||:
    1776 GCAGCCCATCTTCGATGACATGCTGCACATCCTCAACCCCGAGGAGCTCC 1825

     363 rgValIleGluGluIleProGlnAlaGluAspLysLeuAspArgLeuPhe 379
         ::||||||||||||||||||||||||||||||||||||||||||||||||
    1826 ACGTGATCGAGGAGATCCCGCAGGCGGAGGACAAGCTGGACAGGCTGTTT 1875

     380 GluIleIleGlyValLysSerGlnGluAlaSerGlnThrLeuLeuAspSe 396
         ||||||   ||||||||||||||||||||||||||||||||||||||||
    1876 GAGATCGCCGGAGTCAAGAGCCAGGAAGCCAGCCAAACCCTCCTGGACTC 1925

     396 rValTyrSerHisLeuProAspLeuLeu 405
         |||||||||||||||||||||||||||
    1926 TGTCTATAGCCACCTTCCTGACTTACTG 1953

seq_name: gb_pat:AR134340

seq_documentation_block:
LOCUS       AR134340     1815 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 24 from patent US 6194151.
ACCESSION   AR134340
VERSION     AR134340.1  GI:14123245
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1815)
  AUTHORS   Busfield,S.J.
```

```
  TITLE     Molecules of the TNF receptor superfamily and uses therefor
  JOURNAL   Patent: US 6194151-A 24 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .1815
                     /organism="unknown"
BASE COUNT      448 a    529 c    462 g    376 t
ORIGIN


alignment_scores:
           Quality: 1527.00            Length:    294
             Ratio:  5.194              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AR134340   ..

 Align seg 1/1  to: AR134340  from: 1  to: 1815

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 800

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     801 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 850

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     851 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 900

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     901 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 950

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     951 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1000

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1001 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1050

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1051 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1100

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1101 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1150

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1151 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1200

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1201 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1250

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1251 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1300

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1301 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1350

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1351 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1400

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1401 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1450

      234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1451 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 1500

      251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1501 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 1550

      267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
          ||||||||||||||||||||||||||||||||||||||||||||||||||
    1551 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 1600

      284 laLeuLeuThrValGluProSerProGlnAsp 294
          |||||||||||||||||||||||||||||||
    1601 CTCTCCTGACGGTGGAGCCTTCCCCACAGGAT 1632

seq_name: gb_pat:AR134339

seq_documentation_block:
LOCUS       AR134339     2638 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 22 from patent US 6194151.
ACCESSION   AR134339
VERSION     AR134339.1  GI:14123244
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2638)
  AUTHORS   Busfield,S.J.
  TITLE     Molecules of the TNF receptor superfamily and uses therefor
  JOURNAL   Patent: US 6194151-A 22 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .2638
                     /organism="unknown"
BASE COUNT      625 a    749 c    701 g    563 t
ORIGIN


alignment_scores:
           Quality: 1527.00            Length:    294
             Ratio:  5.194              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-733-956-2 x AR134339   ..

 Align seg 1/1  to: AR134339  from: 1  to: 2638

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1260 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 1309

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1310 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 1359

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1360 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 1409

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1410 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 1459

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1460 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 1509

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1510 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 1559
```

```
     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1560 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 1609

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1610 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 1659

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1660 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 1709

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1710 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGA 1759

     167 pIleLeuLysLeuValAlaAlaGlnValGlySerGlnTrpLysAspIleT 184
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1760 TATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCT 1809

     184 yrGlnPheLeuCysAsnAlaSerGluArgGluValAlaAlaPheSerAsn 200
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1810 ATCAGTTTCTTTGCAATGCCAGTGAGAGGGAGGTTGCTGCTTTCTCCAAT 1859

     201 GlyTyrThrAlaAspHisGluArgAlaTyrAlaAlaLeuGlnHisTrpTh 217
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1860 GGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGAC 1909

     217 rIleArgGlyProGluAlaSerLeuAlaGlnLeuIleSerAlaLeuArgG 234
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1910 CATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCC 1959

     234 lnHisArgArgAsnAspValValGluLysIleArgGlyLeuMetGluAsp 250
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1960 AGCACCGGAGAAACGATGTTGTGGAGAAGATTCGTGGGCTGATGGAAGAC 2009

     251 ThrThrGlnLeuGluThrAspLysLeuAlaLeuProMetSerProSerPr 267
         ||||||||||||||||||||||||||||||||||||||||||||||||
    2010 ACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCC 2059

     267 oLeuSerProSerProIleProSerProAsnAlaLysLeuGluAsnSerA 284
         ||||||||||||||||||||||||||||||||||||||||||||||||
    2060 GCTTAGCCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCG 2109

     284 laLeuLeuThrValGluProSerProGlnAsp 294
         |||||||||||||||||||||||||||||||
    2110 CTCTCCTGACGGTGGAGCCTTCCCCACAGGAT 2141

seq_name: gb_pat:AR134328

seq_documentation_block:
LOCUS       AR134328     2612 bp    DNA             PAT       16-MAY-2001
DEFINITION  Sequence 3 from patent US 6194151.
ACCESSION   AR134328
VERSION     AR134328.1  GI:14123233
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2612)
  AUTHORS   Busfield,S.J.
  TITLE     Molecules of the TNF receptor superfamily and uses therefor
  JOURNAL   Patent: US 6194151-A 3 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .2612
                     /organism="unknown"
BASE COUNT      661 a    713 c    650 g    588 t
ORIGIN


alignment_scores:
           Quality: 1269.00            Length:    259
                Ratio:  5.138               Gaps:      1
  Percent Similarity:  95.367   Percent Identity: 94.981

alignment_block:
 US-09-733-956-2 x AR134328   ..

 Align seg 1/1  to: AR134328  from: 1  to: 2612

      11 SerValArgProLysValLeuSerSerIleGlnGluGlyThrValProAs 27
         |||   ||||||||||||||||||||||||||||||||||||||||||
     930 TCCCAAAGACCAAAGGTACTGAGTAGCATCCAGGAAGGGACAGTCCCTGA 979

      27 pAsnThrSerSerAlaArgGlyLysGluAspValAsnLysThrLeuProA 44
         ||||||||||||||||||||||||||||||||||||||||||||||||
     980 CAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGACCCTCCCAA 1029

      44 snLeuGlnValValAsnHisGlnGlnGlyProHisHisArgHisIleLeu 60
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1030 ACCTTCAGGTAGTCAACCACCAGCAAGGCCCCCACCACAGACACATCCTG 1079

      61 LysLeuLeuProSerMetGluAlaThrGlyGlyGluLysSerSerThrPr 77
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1080 AAGCTGCTGCCGTCCATGGAGGCCACTGGGGGCGAGAAGTCCAGCACGCC 1129

      77 oIleLysGlyProLysArgGlyHisProArgGlnAsnLeuHisLysHisP 94
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1130 CATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAGCATT 1179

      94 heAspIleAsnGluHisLeuProTrpMetIleValLeuPheLeuLeuLeu 110
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1180 TTGACATCAATGAGCATTTGCCCTGGATGATTGTGCTTTTCCTGCTGCTG 1229

     111 ValLeuValValIleValValCysSerIleArgLysSerSerArgThrLe 127
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1230 GTGCTTGTGGTGATTGTGGTGTGCAGTATCCGGAAAAGCTCGAGGACTCT 1279

     127 uLysLysGlyProArgGlnAspProSerAlaIleValGluLysAlaGlyL 144
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1280 GAAAAAGGGGCCCCGGCAGGATCCCAGTGCCATTGTGGAAAAGGCAGGGC 1329

     144 euLysLysSerMetThrProThrGlnAsnArgGluLysTrpIleTyrTyr 160
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1330 TGAAGAAATCCATGACTCCAACCCAGAACCGGGAGAAATGGATCTACTAC 1379

     161 CysAsnGlyHisGlyIleAspIleLeuLysLeuValAlaAlaGlnValGl 177
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1380 TGCAATGGCCATGGTATCGATATCCTGAAGCTTGTAGCAGCCCAAGTGGG 1429

     177 ySerGlnTrpLysAspIleTyrGlnPheLeuCysAsnAlaSerGluArgG 194
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1430 AAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAGAGGG 1479

     194 luValAlaAlaPheSerAsnGlyTyrThrAlaAspHisGluArgAlaTyr 210
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1480 AGGTTGCTGCTTTCTCCAATGGGTACACAGCCGACCACGAGCGGGCCTAC 1529

     211 AlaAlaLeuGlnHisTrpThrIleArgGlyProGluAlaSerLeuAlaGl 227
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1530 GCAGCTCTGCAGCACTGGACCATCCGGGGCCCCGAGGCCAGCCTCGCCCA 1579

     227 nLeuIleSerAlaLeuArgGlnHisArgArgAsnAspValValGluLysI 244
         ||||||||||||||||||||||||||||||||||||||||||||||||
    1580 GCTAATTAGCGCCCTGCGCCAGCACCGGAGAAACGATGTTGTGGAGAAGA 1629

     244 leArgGlyLeuMetGluAspThrThrGln..................... 253
         ||||||||||||||||||||||||||||
    1630 TTCGTGGGCTGATGGAAGACACCACCCAGGTAATGGAGCCCTTGTTGTGT 1679

     254 ...LeuGluThrAspLysLeuAlaLeu 261
            :::   ||||||   |||   |||
    1680 GTCATTACCACCGACCTATTGCCCCTA 1706

seq_name: gb_pr:HSJ181J13
```

```
seq_documentation_block:
LOCUS       HSJ181J13  118271 bp    DNA             PRI       21-FEB-2000
DEFINITION  Human DNA sequence from clone RP1-181J13 on chromosome 6p12.2-21.1
            Contains gene for DR6 (TNFR-related death receptor-6), ESTS, STSs,
            GSSs and a CpG Island, complete sequence.
ACCESSION   AL096801
VERSION     AL096801.18  GI:6478163
KEYWORDS    HTG; CpG Island; DR6; TNFR.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 118271)
  AUTHORS   Parker,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 29, 1999 this sequence version replaced gi:6469421.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP1-181J13 is
            from the library RPCI-1 constructed at the Roswell Park Cancer
            Institute by the group of Pieter de Jong. For further details see
            http://bacpac.med.buffalo.edu/
            VECTOR: pCYPAC2
            IMPORTANT: This sequence is not the entire insert of clone
            RP1-181J13 It may be shorter because we only sequence overlapping
            sections once, or longer because we arrange for a small overlap
            between neighbouring submissions.
            The true left end of clone RP1-181J13 is at 1 in this sequence. The
            true left end of clone RP1-1C16 is at 118172 in this sequence. This
            sequence was generated from part of bacterial clone contigs of
            human chromosome 6, constructed by the Sanger Centre Chromosome 6
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6.
FEATURES             Location/Qualifiers
     source          1. .118271
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p12.2-21.1"
                     /clone="RP1-181J13"
                     /clone_lib="RPCI-1"
     repeat_region   4. .215
                     /note="AluSc repeat: matches 93. .303 of consensus"
     repeat_region   1668. .1735
                     /note="17 copies 4 mer tcca 91% conserved"
     repeat_region   1737. .2216
                     /note="MER31B repeat: matches 1. .485 of consensus"
     repeat_region   2223. .2350
                     /note="L2 repeat: matches 2582. .2710 of consensus"
     repeat_region   2370. .2567
                     /note="L2 repeat: matches 2244. .2452 of consensus"
     misc_feature    complement(2700..3156)
                     /note="match: GSS: Em:AQ137790"
     misc_feature    complement(2712. .3043)
                     /note="match: GSS: Em:AQ094362"
     misc_feature    complement(3115. .3466)
                     /note="match: STS: Em:Z53711
                     match: STS: Em:HSB336ZH9"
     repeat_region   3304. .3395
                     /note="46 copies 2 mer tg 73% conserved"
     repeat_region   3305. .3396
                     /note="23 copies 4 mer gtgt 73% conserved"
     repeat_region   3307. .3378
                     /note="3 copies 24 mer 80% conserved"
     misc_feature    3581. .4013
                     /note="match: GSS: Em:AQ030912"
     repeat_region   5096. .5216
                     /note="L1P3 repeat: matches 5617. .5742 of consensus"
     repeat_region   5394. .5677
                     /note="AluSx repeat: matches 15. .300 of consensus"
     repeat_region   5785. .6049
                     /note="L2 repeat: matches 2182. .2455 of consensus"
     repeat_region   6052. .6085
                     /note="17 copies 2 mer tt 85% conserved"
     repeat_region   6089. .7315
                     /note="L1PB1 repeat: matches 4911. .6152 of consensus"
     repeat_region   7326. .7384
                     /note="L2 repeat: matches 2682. .2737 of consensus"
     repeat_region   7525. .7588
                     /note="MLT1D repeat: matches 1. .65 of consensus"
     repeat_region   7767. .7902
                     /note="MLT1D repeat: matches 45. .177 of consensus"
     repeat_region   7963. .8192
                     /note="MIR repeat: matches 13. .262 of consensus"
     repeat_region   8467. .8618
                     /note="MIR repeat: matches 107. .259 of consensus"
     repeat_region   8769. .8812
                     /note="11 copies 4 mer tctc 95% conserved"
     repeat_region   8813. .9282
                     /note="MLT1D repeat: matches 1. .456 of consensus"
     repeat_region   9285. .9548
                     /note="L2 repeat: matches 431. .712 of consensus"
     misc_feature    9549. .9837
                     /note="match: GSS: Em:AQ041048"
     misc_feature    9549. .9853
                     /note="match: GSS: Em:B35323"
     repeat_region   9875. .10219
                     /note="L2 repeat: matches 2241. .2624 of consensus"
     repeat_region   10464. .10534
                     /note="MER21B repeat: matches 12. .81 of consensus"
     repeat_region   10535. .10963
                     /note="MER57A repeat: matches 1. .434 of consensus"
     repeat_region   10964. .11651
                     /note="MER21B repeat: matches 81. .778 of consensus"
     repeat_region   11647. .11717
                     /note="MER21B repeat: matches 724. .794 of consensus"
     repeat_region   12489. .18627
                     /note="L1PA2 repeat: matches 11. .6144 of consensus"
     repeat_region   18671. .18981
                     /note="AluJo repeat: matches 1. .310 of consensus"
     repeat_region   19040. .19169
                     /note="AluJo repeat: matches 7. .135 of consensus"
     repeat_region   19322. .21409
                     /note="L1MA8 repeat: matches 4105. .6230 of consensus"
     repeat_region   21410. .21717
                     /note="AluY repeat: matches 1. .308 of consensus"
     repeat_region   21718. .22040
                     /note="L1MA8 repeat: matches 3791. .4105 of consensus"
     repeat_region   22258. .22496
                     /note="MIR repeat: matches 9. .253 of consensus"
     misc_feature    complement(22497. .22799)
                     /note="match: GSS: Em:AQ548159"
     misc_feature    22811. .23243
                     /note="match: GSS: Em:AQ711306"
     repeat_region   23092. .23152
                     /note="L2 repeat: matches 2688. .2750 of consensus"
     repeat_region   23267. .23306
                     /note="L2 repeat: matches 2671. .2709 of consensus"
```

```
     repeat_region   23473. .23690
                     /note="L2 repeat: matches 2064. .2298 of consensus"
     repeat_region   23900. .23957
                     /note="29 copies 2 mer aa 75% conserved"
     repeat_region   24021. .24312
                     /note="AluSp repeat: matches 1. .294 of consensus"
     repeat_region   24901. .25210
                     /note="AluSg repeat: matches 1. .303 of consensus"
     repeat_region   25394. .25691
                     /note="AluSx repeat: matches 1. .298 of consensus"
     repeat_region   26282. .26417
                     /note="L2 repeat: matches 2570. .2704 of consensus"
     repeat_region   26759. .26842
                     /note="MIR repeat: matches 37. .120 of consensus"
     repeat_region   27026. .27226
                     /note="MIR repeat: matches 35. .249 of consensus"
     misc_feature    27372. .27965
                     /note="match: GSS: Em:AQ240568"
     repeat_region   27966. .28231
                     /note="L2 repeat: matches 2457. .2745 of consensus"
     repeat_region   28497. .28695
                     /note="L2 repeat: matches 2041. .2227 of consensus"
     repeat_region   28784. .28974
                     /note="MLT1J repeat: matches 61. .221 of consensus"
     repeat_region   29201. .29320
                     /note="5 copies 24 mer 72% conserved"
     repeat_region   29289. .29324
                     /note="18 copies 2 mer ct 80% conserved"
     repeat_region   29292. .29323
                     /note="8 copies 4 mer tctc 87% conserved"
     repeat_region   29325. .29610
                     /note="AluSg1 repeat: matches 1. .309 of consensus"
     repeat_region   29693. .30204
                     /note="MER9 repeat: matches 1. .511 of consensus"
     repeat_region   30205. .30768
                     /note="HERVK_9 repeat: matches 1. .568 of consensus"
     repeat_region   30770. .32408
                     /note="HERVK_9 repeat: matches 1113. .2772 of consensus"
     repeat_region   32416. .32572
                     /note="HERVK_9 repeat: matches 3711. .3867 of consensus"
     repeat_region   32659. .33272
                     /note="HERVK_9 repeat: matches 5303. .5916 of consensus"
     repeat_region   33273. .33776
                     /note="MER9 repeat: matches 1. .503 of consensus"
     repeat_region   33964. .34050
                     /note="AluYb8 repeat: matches 231. .317 of consensus"
     misc_feature    34721. .35356
                     /note="match: GSS: Em:AQ265438"
     repeat_region   35982. .36176
                     /note="MER5B repeat: matches 12. .175 of consensus"
     misc_feature    36177. .36582
                     /note="match: STS: Em:HS124B7T"
     misc_feature    36381. .37748
                     /note="CpG island"
                     /evidence=not_experimental
     misc_feature    join(36746. .37052,59873. .60016)

alignment_scores:
            Quality:  874.00              Length:    215
              Ratio:  4.724                 Gaps:      2
 Percent Similarity:  86.047   Percent Identity: 80.000

alignment_block:
 US-09-733-956-2 x HSJ181J13   ..

 Align seg 1/1  to: HSJ181J13  from: 1  to: 118271

      1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62038 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 62087

     17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||

  62088 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 62137

     34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62138 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 62187

     51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62188 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 62237

     67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62238 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 62287

     84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62288 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 62337

    101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62338 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 62387

    117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62388 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 62437

    134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62438 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 62487

    151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  62488 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTGAGCC 62537

    167 pIleLeuLysLeuValAlaAlaGlnVal..................... 176
           |||         ::::::|||:::
  62538 TTCTCTCTTCTCTCCTTGCTCACAGATCTCTCTACCACTTTTTCCCTAGT 62587

    177 .............GlySerGlnTrpLysAspIleTyrGlnPheLeuCys 188
                      |||         ::::::::::::::::::||||||
  62588 CCCTGTTTTCTGAAGGAATGGGAATGCGTAGTTTGTGGAATCTCCTTTGT 62637

    189 AsnAlaSer.......................GluArgGluVal 195
        :::   :::                      |||:::|||:::
  62638 AGACCAAACACTACTTGATCTAGTCGTGTTATAGAAAAAGAGATT 62682

seq_name: gb_htg:AL512350

seq_documentation_block:
LOCUS       AL512350   150180 bp    DNA             HTG       10-JUL-2001
DEFINITION  Homo sapiens chromosome 6 clone RP11-150D3, *** SEQUENCING IN
            PROGRESS ***, 2 unordered pieces.
ACCESSION   AL512350
VERSION     AL512350.7  GI:13169571
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 150180)
  AUTHORS   Corby,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 28, 2001 this sequence version replaced gi:13161707.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA150D3
```

```
            ---------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: M13; M77815; 13% of reads
            Sequencing vector: plasmid; L08752; 86% of reads
            Chemistry: Dye-terminator ABI; 2% of reads
            Chemistry: Dye-terminator Big Dye; 89% of reads
            Chemistry: Dye-primer-amersham; 6% of reads
            Chemistry: Dye-primer Big Dye; 1% of reads
            Consensus quality: 148953 bases at least Q40
            Consensus quality: 149510 bases at least Q30
            Consensus quality: 149849 bases at least Q20
            Insert size: 150080; sum-of-contigs
            Insert size: 143285; 6.0% error; agarose-fp
            Quality coverage: 11.30x in Q20 bases; sum-of-contigs Quality
            coverage: 14.96x in Q20 bases; agarose-fp
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   148067: contig of 148067 bp in length
            *   148068 148167: gap of      100 bp
            *   148168   150180: contig of 2013 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150180
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-150D3"
                     /clone_lib="RPCI-11.1"
     misc_feature    1. .148067
                     /note="assembly_fragment:01601
                     clone_end:SP6
                     vector_side:left"
     misc_feature    148168. .150180
                     /note="assembly_fragment:06349"
BASE COUNT    46523 a  31441 c  29592 g  42524 t    100 others
ORIGIN


alignment_scores:
           Quality:  874.00            Length:    215
             Ratio:  4.724               Gaps:      2
 Percent Similarity:  86.047   Percent Identity: 80.000

alignment_block:
 US-09-733-956-2 x AL512350/rev   ..

 Align seg 1/1  to reverse of: AL512350  from: 1  to: 150180

       1 MetAsnSerThrGluSerAsnSerSerAlaSerValArgProLysValLe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141678 ATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACT 141629

      17 uSerSerIleGlnGluGlyThrValProAspAsnThrSerSerAlaArgG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141628 GAGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGG 141579

      34 lyLysGluAspValAsnLysThrLeuProAsnLeuGlnValValAsnHis 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141578 GGAAGGAAGACGTGAACAAGACCCTCCCAAACCTTCAGGTAGTCAACCAC 141529

      51 GlnGlnGlyProHisHisArgHisIleLeuLysLeuLeuProSerMetGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141528 CAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGA 141479

      67 uAlaThrGlyGlyGluLysSerSerThrProIleLysGlyProLysArgG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141478 GGCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGG 141429

      84 lyHisProArgGlnAsnLeuHisLysHisPheAspIleAsnGluHisLeu 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141428 GACATCCTAGACAGAACCTACACAAGCATTTTGACATCAATGAGCATTTG 141379

     101 ProTrpMetIleValLeuPheLeuLeuLeuValLeuValValIleValVa 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141378 CCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGT 141329

     117 lCysSerIleArgLysSerSerArgThrLeuLysLysGlyProArgGlnA 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141328 GTGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGG 141279

     134 spProSerAlaIleValGluLysAlaGlyLeuLysLysSerMetThrPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141278 ATCCCAGTGCCATTGTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCA 141229

     151 ThrGlnAsnArgGluLysTrpIleTyrTyrCysAsnGlyHisGlyIleAs 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
  141228 ACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTGAGCC 141179

     167 pIleLeuLysLeuValAlaAlaGlnVal..................... 176
             |||         ::::::|||:::
  141178 TTCTCTCTTCTCTCCTTGCTCACAGATCTCTCTACCACTTTTTCCCTAGT 141129

     177 .............GlySerGlnTrpLysAspIleTyrGlnPheLeuCys 188
                       |||        ::::::::::::::::::|||||
  141128 CCCTGTTTTCTGAAGGAATGGGAATGCGTAGTTTGTGGAATCTCCTTTGT 141079

     189 AsnAlaSer.....................GluArgGluVal 195
         :::   :::                    |||:::|||:::
  141078 AGACCAAACACTACTTGATCTAGTCGTGTTATAGAAAAAGAGATT 141034

seq_name: gb_htg:AC026260

seq_documentation_block:
LOCUS       AC026260   125419 bp    DNA             HTG       01-JUN-2000
DEFINITION  Homo sapiens chromosome 1 clone RP11-304O4 map 1, LOW-PASS SEQUENCE
            SAMPLING.
ACCESSION   AC026260
VERSION     AC026260.2  GI:8139093
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 125419)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 1, clone RP11-304O4
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 125419)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
```

```
  JOURNAL  Submitted (21-MAR-2000) Whitehead Institute/MIT Center for Genome
           Research, 320 Charles Street, Cambridge, MA 02141, USA
  COMMENT  On Jun 1, 2000 this sequence version replaced gi:7272272.
           All repeats were identified using RepeatMasker:
           Smit, A.F.A. & Green, P. (1996-1997)
           http://ftp.genome.washington.edu/RM/RepeatMasker.html
           -------------- Genome Center
               Center: Whitehead Institute/ MIT Center for Genome Research
               Center code: WIBR
               Web site: http://www-seq.wi.mit.edu
               Contact: sequence_submissions@genome.wi.mit.edu
           -------------- Project Information
               Center project name: L8128
               Center clone name: 304_O_4
           --------------.
           * NOTE: This record contains 151 individual
           * sequencing reads that have not been assembled into
           * contigs. Runs of N are used to separate the reads
           * and the order in which they appear is completely
           * arbitrary. Low-pass sequence sampling is useful for
           * identifying clones that may be gene-rich and allows
           * overlap relationships among clones to be deduced.
           * However, it should not be assumed that this clone
           * will be sequenced to completion. In the event that
           * the record is updated, the accession number will
           * be preserved.
           *        1      765: contig of 765 bp in length
           *      766 865: gap of      100 bp
           *      866     1613: contig of 748 bp in length
           *     1614 1713: gap of      100 bp
           *     1714     2478: contig of 765 bp in length
           *     2479 2578: gap of      100 bp
           *     2579     3318: contig of 740 bp in length
           *     3319 3418: gap of      100 bp
           *     3419     4158: contig of 740 bp in length
           *     4159 4258: gap of      100 bp
           *     4259     5016: contig of 758 bp in length
           *     5017 5116: gap of      100 bp
           *     5117     5898: contig of 782 bp in length
           *     5899 5998: gap of      100 bp
           *     5999     6729: contig of 731 bp in length
           *     6730 6829: gap of      100 bp
           *     6830     7607: contig of 778 bp in length
           *     7608 7707: gap of      100 bp
           *     7708     8462: contig of 755 bp in length
           *     8463 8562: gap of      100 bp
           *     8563     9304: contig of 742 bp in length
           *     9305 9404: gap of      100 bp
           *     9405    10153: contig of 749 bp in length
           *    10154 10253: gap of      100 bp
           *    10254    11009: contig of 756 bp in length
           *    11010 11109: gap of      100 bp
           *    11110    11845: contig of 736 bp in length
           *    11846 11945: gap of      100 bp
           *    11946    12729: contig of 784 bp in length
           *    12730 12829: gap of      100 bp
           *    12830    13572: contig of 743 bp in length
           *    13573 13672: gap of      100 bp
           *    13673    14397: contig of 725 bp in length
           *    14398 14497: gap of      100 bp
           *    14498    15240: contig of 743 bp in length
           *    15241 15340: gap of      100 bp
           *    15341    16086: contig of 746 bp in length
           *    16087 16186: gap of      100 bp
           *    16187    16907: contig of 721 bp in length
           *    16908 17007: gap of      100 bp
           *    17008    17806: contig of 799 bp in length
           *    17807 17906: gap of      100 bp
           *    17907    18678: contig of 772 bp in length
           *    18679 18778: gap of      100 bp
           *    18779    19523: contig of 745 bp in length
           *    19524 19623: gap of      100 bp
           *    19624    20368: contig of 745 bp in length
           *    20369 20468: gap of      100 bp
           *    20469    21217: contig of 749 bp in length
           *    21218 21317: gap of      100 bp
           *    21318    22052: contig of 735 bp in length
           *    22053 22152: gap of      100 bp
           *    22153    22920: contig of 768 bp in length
           *    22921 23020: gap of      100 bp
           *    23021    23776: contig of 756 bp in length
           *    23777 23876: gap of      100 bp
           *    23877    24606: contig of 730 bp in length
           *    24607 24706: gap of      100 bp
           *    24707    25456: contig of 750 bp in length
           *    25457 25556: gap of      100 bp
           *    25557    26333: contig of 777 bp in length
           *    26334 26433: gap of      100 bp
           *    26434    27207: contig of 774 bp in length
           *    27208 27307: gap of      100 bp
           *    27308    28037: contig of 730 bp in length
           *    28038 28137: gap of      100 bp
           *    28138    28919: contig of 782 bp in length
           *    28920 29019: gap of      100 bp
           *    29020    29771: contig of 752 bp in length
           *    29772 29871: gap of      100 bp
           *    29872    30649: contig of 778 bp in length
           *    30650 30749: gap of      100 bp
           *    30750    31545: contig of 796 bp in length
           *    31546 31645: gap of      100 bp
           *    31646    32375: contig of 730 bp in length
           *    32376 32475: gap of      100 bp
           *    32476    33237: contig of 762 bp in length
           *    33238 33337: gap of      100 bp
           *    33338    34068: contig of 731 bp in length
           *    34069 34168: gap of      100 bp
           *    34169    34911: contig of 743 bp in length
           *    34912 35011: gap of      100 bp
           *    35012    35741: contig of 730 bp in length
           *    35742 35841: gap of      100 bp
           *    35842    36621: contig of 780 bp in length
           *    36622 36721: gap of      100 bp
           *    36722    37482: contig of 761 bp in length
           *    37483 37582: gap of      100 bp
           *    37583    38332: contig of 750 bp in length
           *    38333 38432: gap of      100 bp
           *    38433    39210: contig of 778 bp in length
           *    39211 39310: gap of      100 bp
           *    39311    40095: contig of 785 bp in length
           *    40096 40195: gap of      100 bp
           *    40196    40978: contig of 783 bp in length
           *    40979 41078: gap of      100 bp
           *    41079    41831: contig of 753 bp in length
           *    41832 41931: gap of      100 bp
           *    41932    42673: contig of 742 bp in length
           *    42674 42773: gap of      100 bp
           *    42774    43516: contig of 743 bp in length
           *    43517 43616: gap of      100 bp
           *    43617    44358: contig of 742 bp in length
           *    44359 44458: gap of      100 bp
           *    44459    45217: contig of 759 bp in length
           *    45218 45317: gap of      100 bp
           *    45318    46067: contig of 750 bp in length
           *    46068 46167: gap of      100 bp
           *    46168    46907: contig of 740 bp in length
           *    46908 47007: gap of      100 bp
           *    47008    47735: contig of 728 bp in length
           *    47736 47835: gap of      100 bp
           *    47836    48605: contig of 770 bp in length
           *    48606 48705: gap of      100 bp
           *    48706    49485: contig of 780 bp in length
           *    49486 49585: gap of      100 bp
           *    49586    50330: contig of 745 bp in length
           *    50331 50430: gap of      100 bp
           *    50431    51176: contig of 746 bp in length
           *    51177 51276: gap of      100 bp
```

```
            *    51277    52010: contig of 734 bp in length
            *    52011 52110: gap of      100 bp
            *    52111    52868: contig of 758 bp in length
            *    52869 52968: gap of      100 bp
            *    52969    53704: contig of 736 bp in length
            *    53705 53804: gap of      100 bp
            *    53805    54496: contig of 692 bp in length
            *    54497 54596: gap of      100 bp
            *    54597    55321: contig of 725 bp in length
            *    55322 55421: gap of      100 bp
            *    55422    56136: contig of 715 bp in length
            *    56137 56236: gap of      100 bp
            *    56237    56945: contig of 709 bp in length
            *    56946 57045: gap of      100 bp
            *    57046    57752: contig of 707 bp in length
            *    57753 57852: gap of      100 bp
            *    57853    58556: contig of 704 bp in length

alignment_scores:
           Quality:  408.00              Length:    127
             Ratio:  4.295                 Gaps:      2
 Percent Similarity:  74.803   Percent Identity: 64.567

alignment_block:
 US-09-733-956-2 x AC026260/rev   ..

 Align seg 1/1  to reverse of: AC026260  from: 1  to: 125419

      89 AsnLeuHisLysHisPheAspIleAsnGluHisLeuProTrpMetIleVa 105
         |||   ||||||||||||||||||||||||||||||||||||||||||||
   69948 AACNTACACAAGCATTTTGACATCAATGAGCATTTGCCCTGGATGATTGT 69899

     105 lLeuPheLeuLeuLeuValLeuValValIleValValCysSerIleArgL 122
         |||||||||||||||||||||||||||   ||||||||||||||||||
   69898 GCTTTTCCTGCTGCTGGTGCTTGTGGTGANTGTGGTGTGCAGTATCCGGA 69849

     122 ysSerSerArgThrLeuLysLysGlyProArgGlnAspProSerAlaIle 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   69848 AAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGGATCCCAGTGCCATT 69799

     139 ValGluLysAlaGlyLeuLysLysSerMetThrProThrGlnAsnArgGl 155
         ||||||||||||||||||||||||||||||||||||||||||||||||||
   69798 GTGGAAAAGGCAGGGCTGAAGAAATCCATGACTCCAACCCAGAACCGGGA 69749

     155 uLysTrpIleTyrTyrCysAsnGlyHisGlyIleAspIleLeuLysLeuV 172
         |||||||||||||||||||||||||||||||          |||
   69748 GAAATGGATCTACTACTGCAATGGCCATGGTGAGCCTTCTCTCTTCTCTC 69699

     172 alAlaAlaGlnVal..................................... 176
           ::::::|||:::
   69698 CTTGCTCACAGATCTCTCTACCACTTTTTCCCTAGTCCCTGTTTTCTGAA 69649

     177 GlySerGlnTrpLysAspIleTyrGlnPheLeuCysAsnAlaSer..... 191
         |||         ::::::::::::::::::|||||||:::   :::
   69648 GGAATGGGAATGCGTAGTTTGTGGAATCTCCTTTGTAGACCAAACACTAC 69599

     192 ..................GluArgGluVal 195
                           |||:::|||:::
   69598 TTGATCTAGTCGTGTTATAGAAAAAGAGATT 69568
```

THIS PAGE BLANK (USPTO)